

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2001, 08:52:56 ; Search time 21.65 seconds
(without alignments)
266.888 Million cell updates/sec

Title: US-09-497-591-1
Perfect score: 200
Sequence: 1 ANSFLXXLRHSLRXRCIXX.....XXAKXIFQNVDTLAFWSKH 44

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_TREMBL_16:.*
2: sp_archaea:.*
3: sp_bacteria:.*
4: sp_fungi:.*
5: sp_human:.*
6: sp_invertebrate:.*
7: sp_mhc:.*
8: sp_mammal:.*
9: sp_organelle:.*
10: sp_phage:.*
11: sp_plant:.*
12: sp_protent:.*
13: sp_unclassified:.*
14: sp_vertebrate:.*
15: sp_virus:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158	79.0	456	6	O9TTR0
2	130	65.0	459	6	O9GLP2
3	105	52.5	482	11	O63207
4	93	46.5	444	4	O14339
5	91	45.5	481	11	O54740
6	91	45.5	481	11	O88947
7	89	44.5	218	4	O14668
8	85	42.5	469	6	O9GMD9
9	82	41.0	650	4	O16519
10	82	41.0	650	4	O9NSD0
11	80	40.0	100	4	O15253
12	80	40.0	446	11	O61109
13	79	39.5	138	6	O28994
14	77	38.5	456	6	O14316
15	75	37.5	648	6	O29094
16	74	37.0	607	13	O91001
17	68	34.0	608	13	O9PTW7
18	65.5	32.8	422	4	O15213
19	64	32.0	25	11	O9OVH6

Result No.	Score	Query Match	Length	ID	Description
20	63	31.5	98	13	P82807
21	62	31.0	673	11	O61592
22	61	30.5	674	11	O63772
23	60	30.0	678	4	O14393
24	56	28.0	202	4	O14669
25	52	26.0	724	4	O9H6S7
26	51.5	25.8	594	10	O9S1G9
27	51	25.5	16	4	O9UC55
28	50.5	25.2	459	10	O9SE22
29	50	25.0	615	10	O9LU67
30	49.5	24.8	510	2	O9POS4
31	49.5	24.8	603	10	O9LPG7
32	49.5	24.8	608	10	O9XFP6
33	48.5	24.2	196	10	O04284
34	48.5	24.2	593	10	O9LUC3
35	48.5	24.2	2454	3	O9UVP2
36	48.5	24.2	2454	3	O9UVS6
37	48	24.0	606	3	O9US07
38	47.5	23.8	506	10	O9SPR0
39	47.5	23.8	543	10	O9SE23
40	47.5	23.8	589	10	O9MB23
41	47.5	23.8	221	6	O9LMS2
42	47	23.5	390	4	O9UG06
43	47	23.5	538	4	O9Y4C9
44	47	23.5	574	4	O9H1G5
45	47	23.5	574	4	O9H1G5

ALIGNMENTS

RESULT 1

ID	Query Match	Length	ID	Description
O9TTR0	79.0	456	6	O9TTR0
AC	O9TTR0	PRELIMINARY	PRT	456 AA.
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	PROTEIN C PRECURSOR.			
GN	PROC.			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Leeb T., Kopp T., Deppe A., Breen M., Matis U., Brunberg L.,			
RA	Brenig B.;			
RT	"Molecular characterization and chromosomal assignment of the canine			
RT	protein C gene.";			
RL	Mamm. Genome 10:135-139(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99371952; PubMed=10443005;			
RA	Leeb T., Pfeiffer I., Kopp T., Deppe A., Brenig B.;			
RT	"Analysis of canine protein C gene polymorphisms.";			
RL	Anim. Genet. 30:237-238(1999).			
CC	-1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.			
CC	-1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).			
DR	EMBL; AJ001979; CAA05126.1; -			
DR	HSSP; P04070; IPCU.			
DR	InterPro: IPR000152; -			
DR	InterPro: IPR000294; -			
DR	InterPro: IPR000561; -			
DR	InterPro: IPR001254; -			
DR	InterPro: IPR001314; -			
DR	InterPro: IPR001881; -			
DR	InterPro: IPR002383; -			
DR	Pfam: PF00008; EGF. 2.			
DR	Pfam: PF00089; trypsin. 1.			
DR	Pfam: PF00594; gla. 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00001; GLABLOOD.			

```

DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR SMART; SM00181; EGF; 1.
KM Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;
KW Hydroxylation; Repeat; Serine protease; Signal.
FT SIGNAL; 1
FT CHAIN; 43 192 PROTEIN C LIGHT CHAIN.
FT CHAIN; 193 194 PROTEIN C CONNECTING DIPEPTIDE.
FT CHAIN; 195 456 PROTEIN C HEAVY CHAIN.
SO SEQUENCE 456 AA; 50813 MW; 7AD3A8C1C34E59FE CRC64;

Query Match 79.0%; Score 158; DB 6; Length 456;
Best Local Similarity 68.2%; Pred. No. 5.7e-19;
Matches 30; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 1 ANSFLXXLRHGSLKRCIXXCDFXKXKXIFQNDPDLAFMSKH 44
DB 43 ANSFLEIRAGSLERECMEETCODEEAKETIFQNDPDLAFMSKY 86

RESULT 2
O9GLP2 PRELIMINARY; PRT; 459 AA.
AC O9GLP2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-NOV-1996 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PROTEIN C.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP TISSUE=LIVER;
RA Grimm D.R., Colter M.B., Kim H.;
RT Cloning of the complete cDNA sequences encoding porcine factor V and
RL protein C.
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF191307; AAC28380.1; -
SO SEQUENCE 459 AA; 51866 MW; 8541AAC14CC16D09 CRC64;

Query Match 65.0%; Score 130; DB 6; Length 459;
Best Local Similarity 56.8%; Pred. No. 4e-14;
Matches 25; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 ANSFLXXLRHGSLKRCIXXCDFXKXKXIFQNDPDLAFMSKH 44
DB 42 ANSFLEIRPSSLERECKETCTDEEARLFTQNTMTAFMSKY 85

RESULT 3
O63207 PRELIMINARY; PRT; 482 AA.
AC O63207;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE FACTOR X.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;

```

```

RX MEDLINE=96093366; PubMed=8578539;
RA Stanton C., Ross R.P., Hutson S.;
RT "Evidence for competition between vitamin K-dependent clotting factors
RT for intracellular processing by the vitamin K-dependent gamma-
RT carboxylase.";
RL Thromb. Res. 80:63-73(1995).
CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR EMBL; X79807; CAA56202.1; -
DR HSSP; P00742; 1XKA.
DR MEROPS; S01.216; -.
DR Interpro; IPR000152; -.
DR Interpro; IPR000294; -.
DR Interpro; IPR000561; -.
DR Interpro; IPR000742; -.
DR Interpro; IPR001254; -.
DR Interpro; IPR001314; -.
DR Interpro; IPR001881; -.
DR Interpro; IPR002383; -.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00001; GLABLOOD.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR SMART; SM00020; TRYP_SPE; 1.
KM Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase;
KW Hydroxylation; Repeat; Serine protease.
SO SEQUENCE 482 AA; 54265 MW; 0284678B3954A698 CRC64;

Query Match 52.5%; Score 105; DB 11; Length 482;
Best Local Similarity 38.6%; Pred. No. 8.9e-10;
Matches 17; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

OY 1 ANSFLXXLRHGSLKRCIXXCDFXKXKXIFQNDPDLAFMSKH 44
DB 41 ANSFEEITKGNLERECVEETCSFEAREVEEDNEKTEFWNKY 84

RESULT 4
O14339 PRELIMINARY; PRT; 444 AA.
AC O14339;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PREPROFACTOR VIIb PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA MEDLINE=86205965; PubMed=3486420;
RA Hagen F.S., Gray C.L., O'Hara P.J., Grant F.J., Saart G.C.,
RA Woodbury R.G., Hart C.E., Insley M., Kistiel W., Kurachi K.,
RA Davie E.W.;
RT "Characterization of a cDNA coding for human factor VII.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:2412-2416(1986).
CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR EMBL; M13232; AAA88041.1; -
DR HSSP; P08709; 1BF9.
DR Interpro; IPR000152; -.
DR Interpro; IPR000294; -.

```


DR	PROSITE: PS01186;	EGF_2;	2.
DR	PROSITE: PS01187;	EGF_CA;	1.
DR	PROSITE: PS00011;	GLU_CARBOXYLATION;	1.
DR	PROSITE: PS00135;	TRYPSIN_HIS;	UNKNOWN_1.
DR	PROSITE: PS00135;	TRYPSIN_SER;	1.
DR	SMART: SM00020;	Tryp_Spc;	1.
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat;		
KW	Serine protease; Signal.		
FT	SIGNAL	1	40
FT	CHAIN	41	481
SO	SEQUENCE	481 AA;	54018 MW; 8AC09DE5EF9D271E CAC64;

Query Match	45.5%	Score 91	DB 11	Length 481
Best Local Similarity	34.1%	Pred. No.	2.3e-07	
Matches 15	Conservative 11	Mismatches 18	Indels 0	Gaps 0

Qy 1 ANSFLXXLRHSGIXRXCIXICDEXXAKXIFOWDDTLAFWSH 44
||| | | | | : || : || : || : || : || : || : || :
Db 41 ANSFEEFKKGNLERECMEETCSYEAEVREIIFDEDEKTRKEYWRY 84

RESULT	7
ID	014668
PRELIMINARY:	
PRT:	218 AA.

DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE PROLINE-RICH GLA PROTEIN 1.

0C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
0C Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
0X NCBI_TaxID=9606;
0X taxid=9606;

RA Kulman J.D., Harris J.E., Haldeman B.A., Davie E.W.,
Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1997).
EMBL: AF009342; AAB67070.1; -.
DR accession: D00746; Genbank: AF009342.

```
DR      Interpro: IPR002383; -.
DR      Pfam: PF00594; glr_1.
DR      PRINTS: PR00001; GLABLOOD.
DR      PROSITE: PS00001; GLABLOOD.OVLTATION. 1
```

Overlaid Match	44.5%	Score: 89	DP: 4	Insert: 318
Sequence	218 AA	24947 MW	26538A61AB0AE98	CRC64

Matches	16;	Conservative	7;	Mismatches	21;	Indels	0;	Gaps	0;
OY	1	ANSELXXLRHGS	SLXRCIXXIC	DFXXAKXII	IQNVDDI	LA	FWSMH	44	

RESULT 8

AC	Q9GMD9,	(Tremblrel, 16, Created)
DT	01-MAR-2001	(Tremblrel, 16, Last sequence update)
DT	01-MAR-2001	(Tremblrel, 16, Last sequence update)
DT	01-MAR-2001	(Tremblrel, 16, Last sequence update)

OS Ornithorhynchus anatinus (Duckbill platypus).
 OS Ornithorhynchus anatinus (Duckbill platypus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.

RP SEQUENCE FROM N. A.

RA Poorafshar M.M., Hellman L.L.:
RT "Identification and structural analysis of three serine proteases in a
RT monotreme, platypus, Ornithorhynchus anatinus.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF275654; AAC00453.1;
SQ SEQUENCE 469 AA; 52196 MW; 4C66C23D00758F6A CRC64;

Query Match	42.5%	Score 85:	DB 6:	length 469;
Best Local Similarity	35.7%	Pred. No.	2.5e-06:	
Matches 15; Conservative		8; Mismatches	19; Indels	0; Gaps 0;

Db 41 ANSLFEELKGNLRECNELTCSYEELAREVEEDTDKINFWN 82

RESULT 9
Q16519
ID Q16519 PRELIMINARY; PRT; 650 AA.

DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
DE PROTEIN S PRECURSOR.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

RX MDLINE=86313649; Pubmed=2944113;
RA Lundwall A., Dackowski W., Cohen E., Shaffer M., Mair A., Dahlback B.,
Sienfio J., Wydro R.;

DR HSSP; P072004.1 EGT.
EMBL; M14338; AAA60181.1; -
RL Proc. Natl. Acad. Sci. U.S.A. 83:6716-6720(1986).
DR HSSP; P072004.1 EGT.

DR	InterPro	IPR0005561	-
DR	InterPro	IPR001791	-
DR	InterPro	IPR001881	-
DR	InterPro	IPR003983	-

```
DR      Pflam; PF00054; laminin_G; 1.
DR      Pflam; PF00594; gla; 1.
DR      PRINTS: PR00001; GLABLOOD.
DR      PROSITE: PS00001; ASV_HYDROXYT; 3
```

DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 2.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR CDD: C000302; EGF_1

KN	Signal.	15	POTENTIAL.
FT	SIGNAL.	16	POTENTIAL.
CC	CHAIN	650	POTENTIAL.
	CHROMAT.	720463	MS. AB8C04AC602BE474 CBG64A.

Query Match	41.0%	Score 82	DB 4	Length 650
Best Local Similarity	36.4%	Pred. NO	1.2e-05	

0y 1 ANSFLXXLRHGSLXRKCIXXICDFXXXAKXIIFNVDDTLAFMSRK 44
 :
 :
 :
Db 16 ANSLLSEETKOGNLERECIEELCNKEAREVFENDPNDPEIDYFYKY 59

RESULT 10

Query Match	40.0%;	Score 80;	DB 11;	Length 446;
Best Local Similarity	43.9%;	Pred. No. 1.7e-05;		
Matches 18; Conservative	4;	Mismatches 19;	Indels 0;	Gaps 0;

```

0Y      1 ANSEFLXXLRHGSLAXRXCIXICDFXXAKXIFQWVDITAFW 41
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     42 ANSLLEELMPGSLERECNEGCSFEAEAEIFKSPERTQFW 82

```

RESULT	13
Q28994	
ID	Q28994
PRELIMINARY;	
PRT;	138 AA

DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE MATURE PORCINE FACTOR IX (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96003866; PubMed=7568220;
RA Brandstetter H., Bauer M., Huber R., Lollar P., Bode W.;
RT "X-ray structure of clotting factor IXa: active site and module
structure related to Xase activity and hemophilia B.";
LT Proc. Natl. Acad. Sci. U.S.A. 92:9796-9800(1995).

RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Iollar P.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBI databases

DR	InterPro: IPR000152;	-	
DR	InterPro: IPR000294;	-	
DR	InterPro: IPR000561;	-	
DR	InterPro: IPR000742;	-	
DR	InterPro: IPR001438;	-	
DR	InterPro: IPR001881;	-	
DR	InterPro: IPR002383;	-	
DR	Pfam: PF00008; EGF_2		
DR	Pfam: PF00594; gla_1		
DR	PRINTS: PR00010; EGFBLD.		
DR	PRINTS: PR00001; GLABLOD.		
DR	PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.		
DR	PROSITE: PS00022; EGF_1; UNKNOWN_1.		
DR	PROSITE: PS01186; EGF_2; 2.		
DR	PROSITE: PS01187; EGF_CA; 1.		
DR	SMART: SM00001; EGF-like; 1.		
KW	Calcium-binding; EGF-like domain; Glycoprotein; Repeat		
FT	NON_TER	1	
FT	NON_TER	138	
SQ	SEQUENCE	138 AA; 15515 MW; 793BABDE4D5FAFAD CRC64;	

Query Match	39.5%;	Score 79;	DB 6;	Length 138;
Best Local Similarity	38.2%;	Pred. No. 7.3e-06;		
Matches 13; Conservative	7;	Mismatches 14;	Indels 0;	Gaps 0;

QY 11 GSLRXRCIXXICDFXXAKXIFQNVDITLAFWSKH 44
| : | | | : : : | : : :
Db 4 GNLERECIEEKCSFEAREVEFENTTEKTNFEWKQY 37

RESULT 14
Q14316
ID Q14316 PRELIMINARY; PRT; 456 AA

DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE F9 (COAGULATION FACTOR IX (PLASMA THROMBOPLASTIC COMPONENT, CHRISTMAS
DE DISEASE, HAEMOPHILIA B)) (FACTOR IX).

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases

RP SEQUENCE OF 3-19 FROM N.A.
RX MEDLINE=88327116; PubMed=3416069;
RA Retsina P.A., Bartina R.M., Ploos van Amstel J.K., Riemens A.
RA Briet E.;
RT "The putative factor IX gene promoter in hemophilia B Leyden
RL Blood 72:1074-1076(1988).
CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1)

```
DR HSPD: P007400, 1C6F, -
DR InterPro: IPR0000152, -
DR InterPro: IPR0000294, -
DR InterPro: IPR000361, -
DR InterPro: IPR000742, -
DR InterPro: IPR001254, -
DR InterPro: IPR001314, -
DR InterPro: IPR001487, -
DR InterPro: IPR001881, -
DR InterPro: IPR002383, -
DR InterPro: IPR002383, -
DR Pfam: PF000089, EGF_2, -
DR Pfam: PF000089, tyrosin_1
```

DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0010; EGFBLD.
DR PRINTS; PRO0001; GLABLOOD.
DR PROSITE; PS00001; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE; PS00134; TRYPsin HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin SER; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Repeat.
KW Serine protease.
QO SEQUENCE 456 AA: 51149 MW: 54E20A1B3964E234 CRC64;

Query Match	38.5%	Score 77;	DB 4;	Length 456;
Best Local Similarity	35.3%	Pred. No. 5.8e-05;		
Matches 12;	Conservative	8;	Mismatches 14;	Indels 0;
			Gaps	0;

```
QY      11 GSLRXRCIXXICDFXXAKXIFQNVDDTLAFWSKH 44
        |:| | : | :| :| :| ::
Db     53 GNLERCEMEKCSFEAREVFENTERTEENKQY 86
```

RESULT	15	
Q29094		
ID	Q29094	PRELIMINARY; PRT; 648 AA

DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE PROTEIN S (FRAGMENT).
GN PROS.
OS Sus scrofa (Pig)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus
XC NCBI_TaxID=9823;

```
RM [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER:
RX MEDLINE-95134217; PubMed-7832752;
RA Greengard J.S., Fernandez J.A., Radtke K.P., Griffin J.H.;
RT *Identification of candidate residues for interaction of protein S
RT with Cab binding protein and activated protein C.*;
RL Blochem. J. 305:397-403(1995).
DR EMBL: L31379; AAA70382.1; -.
DR HSSP: P00740; ICFH.
DR InterPro: IPR00152; -.
DR InterPro: IPR000294; -.
DR InterPro: IPR000561; -.
DR InterPro: IPR001791; -.
DR InterPro: IPR001881; -.
DR InterPro: IPR002383; -.
DR Pfam: PF00008; EGF; 4.
DR Pfam: PF00054; Laminin_G; 1.
DR PRINTS: PR00001; GLABLOOD.
DR PROSITE: PS00010; ASX_HYDROXYL; 4.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 3.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR SMART: SM00282; LamG; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat;
KW Vitamin K.
FT NON_TER 1
SQ SEQUENCE 648 AA; 71914 MW; 4BDF9AB813872136 CRC64;
```

```
Query Match 37.5%; Score 75; DB 6; Length 648;
Best Local Similarity 36.6%; Pred. No. 0.00019;
Matches 15; Conservative 8; Mismatches 18; Indels 0; Gaps 0;
```

```
OY 1 ANSFLXLRHGSILRXICIXICDFFXXAKXIFQNVDDTLAFW 41
   ||| | : | | | | | | | | | | | | | | | |
DB 15 ANSMLEERKQGNLERECIEELCSKEAREVFEVNDPEYEFY 55
```

Search completed: November 8, 2001, 08:55:54
Job time: 178 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2001, 08:53:16 ; Search time 9.93 Seconds

(without alignments)
151.787 Million cell updates/sec

Title: US-09-497-591-1

Perfect score: 200
Sequence: 1 ANSFLXLRHSLRXRCIX.....XXAKXIFQNVDTLAFWSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
	1	176	88.0	461	1	PRTG_HUMAN
	2	147	73.5	461	1	PRTG_MOUSE
	3	146	73.0	461	1	PRTG_RAT
	4	141	70.5	458	1	PRTG_BOVINE
	5	129	64.5	456	1	PRTG_BOVINE
	6	104	52.0	444	1	FA7_RABIT
	7	101	50.5	492	1	FA10_BOVINE
	8	100	50.0	488	1	FA10_HUMAN
	9	93	46.5	466	1	FA7_HUMAN
	10	93	46.5	490	1	FA7_BOVINE
	11	90	45.0	407	1	FA7_BOVINE
	12	82.5	41.2	617	1	THRB_RAT
	13	82.5	41.2	618	1	THRB_MOUSE
	14	82	41.0	649	1	PRTS_MACMU
	15	82	41.0	676	1	PRTS_HUMAN
	16	81	40.5	452	1	FA9_CANFA
	17	81	40.5	459	1	FA9_MOUSE
	18	81	40.5	646	1	PRTS_RABIT
	19	80	40.0	446	1	FA7_MOUSE
	20	80	40.0	622	1	THRB_HUMAN
	21	80	40.0	675	1	PRTS_BOVINE
	22	78	39.0	675	1	PRTS_RAT
	23	77	38.5	461	1	FA9_HUMAN
	24	74	37.0	416	1	FA9_BOVINE
	25	73	36.5	376	1	FA10_TROCA
	26	73	36.5	475	1	FA10_CHICK
	27	71	35.5	675	1	PRTS_MOUSE
	28	68	34.0	625	1	THRB_BOVINE
	29	65.5	32.8	400	1	PRTZ_HUMAN
	30	59.5	29.8	396	1	PRTZ_BOVINE
	31	53	28.5	501	1	MCOL_CANAL
	32	51	25.5	245	1	LUXP_VIBHA
	33	50	25.0	363	1	ADK_TOXGO

34	49	24.5	422	1	SPM1_SCHPO	O92398 schizosacch
35	49	24.5	730	1	PR1A_HAETN	P44647 haemophilus
36	47.5	23.8	489	1	MEP3_YEAST	P53930 saccharomyc
37	47	23.5	467	1	VL2_HPV40	P36760 human papil
38	47	23.5	1333	1	VGRI_MOUSE	P35969 mus musculu
39	47	23.5	1336	1	VGRI_RAT	P17948 ratus norv
40	47	23.5	1338	1	VGRI_HUMAN	P44077 haemophilus
41	46	23.0	206	1	V930_HAETN	P37619 escherichia
42	46	23.0	221	1	YHQO_ECOLI	P04390 escherichia
43	46	23.0	244	1	T2E5_ECOLI	P40260 saccharomyc
44	45.5	22.8	492	1	MEP1_YEAST	O9y2a9 h probable
45	44	22.0	372	1	B3G8_HUMAN	

ALIGNMENTS

RESULT 1					
ID	PRTC_HUMAN	STANDARD:	PRT:	461	AA.
AC	P04070: Q16001; Q15190; Q15189;				
DT	01-NOV-1986 (Rel. 03, Created)				
DT	01-NOV-1986 (Rel. 03, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	VITAMIN-K DEPENDENT PROTEIN C PRECURSOR (EC 3.4.21.69)				
DE	(AUTOPROTHROMBIN IIA) (ANTICOAGULANT PROTEIN C) (BLOOD COAGULATION				
DE	FACTOR XIV).				
GN	PROC.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OX	NCBI_Taxid:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85270390; PubMed=2991887;				
RA	Foster D.C., Yoshitake S., Davie E.W.;				
RT	"The nucleotide sequence of the gene for human protein C.;"				
RL	Proc. Natl. Acad. Sci. U.S.A. 82:4673-4677(1985).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=8526639; PubMed=2991859;				
RA	Beckmann R.J., Schmidt R.J., Santeiro R.F., Plutzky J., Crabtree G.R.,				
RT	Long G.L.;				
RT	"The structure and evolution of a 461 amino acid human protein C				
RT	precursor and its messenger RNA, based upon the DNA sequence of				
RT	cloned human liver cDNAs.;"				
RL	Nucleic Acids Res. 13:5233-5247(1985).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=86120978; PubMed=3511471;				
RA	Plutzky J., Hoskins J.A., Long G.L., Crabtree G.R.;				
RT	"Evolution and organization of the human protein C gene.;"				
RL	Proc. Natl. Acad. Sci. U.S.A. 83:546-550(1986).				
RN	[4]				
RP	SEQUENCE OF 106-461 FROM N.A.				
RX	MEDLINE=84272714; PubMed=6589623;				
RA	Foster D.C., Davie E.W.;				
RT	"Characterization of a cDNA coding for human protein C.;"				
RL	Proc. Natl. Acad. Sci. U.S.A. 81:4766-4770(1984).				
RN	[5]				
RP	CARBOHYDRATE-LINKAGE SITE ASN-371.				
RX	MEDLINE=90293094; PubMed=1694179;				
RA	Milletich J.P., Broze G.J. Jr.;				
RT	"Beta protein C is not glycosylated at asparagine 329. The rate of				
RT	cysteine sites.;"				
RL	J. Biol. Chem. 265:11397-11404(1990).				
RN	[6]				
RP	HYDROXYLATION.				
RX	MEDLINE=92184750; PubMed=1544894;				
RA	Harris R.J., Ling V.T., Spelman M.W.;				
RT	"O-linked fucose is present in the first epidermal growth factor				
RT	domain of factor XII but not protein C.;"				

RL J. Biol. Chem. 267:5102-5107(1992).
RN [7]
RP 3D-STRUCTURE MODELING OF 175-450.
RX MEDLINE=94272342; PubMed=8003977;
RA Fisher C.L., Greengard J.S., Griffin J.H.;
RT "Models of the serine protease domain of the human antithrombotic
RT plasma factor activated protein C and its zymogen.";
RL Protein Sci. 3:588-599(1994).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 84-461.
RX MEDLINE=97157472; PubMed=9003757;
RA Mather T., Ogaessyan V., Hof P., Huber R., Foundling S., Esmon C.,
RA Bode W.;
RT "The 2.8 A crystal structure of Glu-domainless activated protein C.";
RL EMBO J. 15:6822-6831(1996).
RN [9]
RP REVIEW ON PROC VARIANTS.
RX MEDLINE=93190290; PubMed=8446940;
RA Reitsma P.H., Poort S.R., Bernardi F., Gandrille S., Long G.L.,
RA Sala N., Cooper D.N.;
RT "Protein C deficiency: a database of mutations. For the Protein C & S
RT Subcommittee of the Scientific and Standardization Committee of the
RT International Society on Thrombosis and Haemostasis.";
RL Thromb. Haemost. 69:77-84(1993).
RN [10]
RP VARIANT CYS-444.
RX MEDLINE=87204221; PubMed=2437584;
RA Romeo G., Hassan H.J., Staemfli S., Roncuizi L., Cianetti L.,
RA Leonardi A., Vicente V., Mannucci P.M., Bertina R., Peschle C.,
RA Cortese R.;
RT "Hereditary thrombophilia: identification of nonsense and missense
RT mutations in the protein C gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2829-2832(1987).
RN [11]
RP VARIANT TRP-211 (LONDON-1).
RX MEDLINE=90089906; PubMed=2602169;
RA Grundy C., Chittoldi A., Talbot S., Bevan D., Kakkar V., Cooper D.N.;
RT "Protein C London 1: recurrent mutation at Arg-169 (CGG-->TGG) in
RT the protein C gene causing thrombosis.";
RL Nucleic Acids Res. 17:10513-10513(1989).
RN [12]
RP VARIANT CYS-272.
RX MEDLINE=91329836; PubMed=1866249;
RA Reitsma P.H., Poort S.R., Allaart C.F., Briet E., Bertina R.M.;
RT "The spectrum of genetic defects in a panel of 40 Dutch families with
RT symptomatic protein C deficiency type I: heterogeneity and founder
RT effects.";
RL Blood 78:890-894(1991).
RN [13]
RP VARIANTS ALA-62 (VERMONT-1) AND MET-76.
RX MEDLINE=92190481; PubMed=1347706;
RA Bovill E.G., Tomczak J.A., Grant B., Bhushan F., Pillemer E.,
RA Rainville I.R., Long G.L.;
RT "Protein C Vermont: symptomatic type II protein C deficiency
RT associated with two Glu domain mutations.";
RL Blood 79:1456-1465(1992).
RN [14]
RP VARIANT ASP-418 (HONG KONG-2).
RX MEDLINE=92305321; PubMed=1611081;
RA Sugahara Y., Miura O., Yuen P., Aoki N.;
RT "Protein C deficiency Hong Kong 1 and 2: hereditary protein C
RT deficiency caused by two mutant alleles, a 5-nucleotide deletion and
RT a missense mutation.";
RL Blood 80:126-133(1992).
RN [15]
RP VARIANT LEU-289.
RX MEDLINE=92380660; PubMed=1511989;
RA Grundy C.B., Chisholm M., Kakkar V.V., Cooper D.N.;
RT "A novel homozygous missense mutation in the protein C (PROC) gene
RT causing recurrent venous thrombosis.";
RL Hum. Genet. 89:683-684(1992).
RN [16]
RP VARIANTS GLN-220 AND TRP-220.
RX MEDLINE=92380661; PubMed=1511989;
RA Grundy C.B., Schulman S., Tengborn L., Kakkar V.V., Cooper D.N.;
RT "Two different missense mutations at Arg 178 of the protein C (PROC)
RT gene causing recurrent venous thrombosis.";
RL Hum. Genet. 89:685-686(1992).
RN [17]
RP VARIANT GLN-220.
RX MEDLINE=93250852; PubMed=1301959;
RA Gandrille S., Vidaud M., Alach M., Alhenc-Gelas M., Fischer A.M.,
RA Gouault-Heilman M., Roulon P., Flessinger J.N., Goossens M.;
RT "Two novel mutations responsible for hereditary type I protein C
RT deficiency: characterization by denaturing gradient gel
RT electrophoresis.";
RL Hum. Mutat. 1:491-500(1992).
RN [18]
RP VARIANT SER-334.
RX MEDLINE=92276939; PubMed=1593215;
RA Yamamoto K., Matsushita T., Sugitara I., Takamatsu J., Iwaseki E.,
RA Wada H., Deguchi K., Shirakawa S., Saito H.;
RT "Homozygous protein C deficiency: identification of a novel missense
RT mutation that causes impaired secretion of the mutant protein C.";
RL J. Lab. Clin. Med. 119:682-689(1992).
RN [19]
RP VARIANTS TRP-38; CYS-42; HIS-42; GLN-271 AND ASN-294.
RX MEDLINE=93313192; PubMed=8324221;
RA Gandrille S., Alhenc-Gelas M., Gaussem P., Allaud M.-F., Dupuy E.,
RA Jhan-Vague I., Alach M.;
RT "Five novel mutations located in exons III and IX of the protein C
RT gene in patients presenting with defective protein C anticoagulant
RT activity.";
RL Blood 82:159-168(1993).
RN [20]
RP VARIANTS G-14; Q-211; Y-244; Q-253; L-321; C-328; I-385; T-388; V-388.
RX MEDLINE=93271391; PubMed=8499565;
RA Poort S.R., Pabinger-Fasching I., Mannhalter C., Reitsma P.H.,
RA Bertina R.M.;
RT "Twelve novel and two recurrent mutations in 14 Austrian families
RT with hereditary protein C deficiency.";
RL Blood Coagul. Fibrinolysis 4:273-280(1993).
RN [21]
RP VARIANT TRP-57.
RX MEDLINE=93271396; PubMed=8499568;
RA Millar D.S., Grundy C.B., Bignelli P., Moffat E.H., Martin R.,
RA Kakkar V.V., Cooper D.N.;
RT "A Glu domain mutation (Arg 15-->Tyr) in the protein C (PROC) gene
RT causing type 2 protein C deficiency and recurrent venous
RT thrombosis.";
RL Blood Coagul. Fibrinolysis 4:345-347(1993).
RN [22]
RP VARIANTS R-145; L-210; W-211; T-243; L-321; M-340 AND Y-426.
RX MEDLINE=94122329; PubMed=8292730;
RA Tsay W., Greengard J.S., Montgomery R.R., McPherson R.A., Fucci J.C.,
RA Koepfer M.A., Coughlin J., Griffin J.H.;
RT "Genetic mutations in ten unrelated American patients with
RT symptomatic type I protein C deficiency.";
RL Blood Coagul. Fibrinolysis 4:791-796(1993).
RN [23]
RP VARIANT SER-423.
RX MEDLINE=94001606; PubMed=8398832;
RA Marchetti G., Patricechhi P., Gemmati D., Castaman G., Rodeghiero F.,
RA Wecey A., Cooper D.N., Tuddenham E.G., Bernardi F.;
RT "Symptomatic type II protein C deficiency caused by a missense
RT mutation (Gly 381-->Ser) in the substrate-binding pocket.";
RL Br. J. Haematol. 84:285-289(1993).
RN [24]
RP SEQUENCE OF 43-64 FROM N.A., AND VARIANT GLY-57 (YONAGO).
RX MEDLINE=93237511; PubMed=8477066;
RA Mimuro J., Muramatsu S., Kaneko M., Yoshitake S., Iijima K.,
RA Nakamura K., Sakata Y., Matsuda M.;
RT "An abnormal protein C (protein C Yonago) with an amino acid
RT substitution of Gly for Arg-15 caused by a single base mutation of C
RT to G in codon 57 (CGG-->GGG). Deteriorated calcium-dependent
RT conformation of the gamma-carboxyglutamic acid domain relevant to a

[illegible]

AC P00745;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VITAMIN K-DEPENDENT PROTEIN C PRECURSOR (EC 3.4.21.69)
DE (AUTOPROTHROMBIN IIA) (ANTICOAGULANT PROTEIN C) (BLOOD COAGULATION
DE FACTOR XIV) (FRAGMENT).
GN PROC.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85014826; PubMed=6091100;
RA Long G.L., Balagaje R.M., McGillivray R.T.A.;
RT "Cloning and sequencing of liver cDNA coding for bovine protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5653-5656(1984).
RN [2]
RP SEQUENCE OF 40-194.
RX MEDLINE=83007325; PubMed=6896876;
RA Fernlund P., Stenflo J.;
RT "Amino acid sequence of the light chain of bovine protein C.";
RL J. Biol. Chem. 257:12170-12179(1982).
RN [3]
RP REVISION TO 110.
RX MEDLINE=83169769; PubMed=6572939;
RA Drakenberg T., Fernlund P., Roepstorff P., Stenflo J.;
RT "beta-hydroxyaspartic acid in vitamin K-dependent protein C.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:1802-1806(1983).
RN [4]
RP SEQUENCE OF 197-456.
RX MEDLINE=83007326; PubMed=6896877;
RA Stenflo J., Fernlund P.;
RT "Amino acid sequence of the heavy chain of bovine protein C.";
RL J. Biol. Chem. 257:12180-12190(1982).
RN [5]
RP PROCESSING, AND CALCIUM-BINDING, DATA.
RX MEDLINE=83213513; PubMed=6304092;
RA Esmon N.L., Debault L.E., Esmon C.T.;
RT "Proteolytic formation and properties of gamma-carboxyglutamic acid-
RT domainless protein C.";
RL J. Biol. Chem. 258:5548-5553(1983).
RN [6]
RP PROCESSING, AND CALCIUM-BINDING, DATA.
RX MEDLINE=83213514; PubMed=6406503;
RA Johnson A.E., Esmon N.L., Laue T.M., Esmon C.T.;
RT "Structural changes required for activation of protein C are induced
RT by Ca2+ binding to a high affinity site that does not contain gamma-
RT carboxyglutamic acid.";
RL J. Biol. Chem. 258:5554-5560(1983).
CC - FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.
CC - CATALYTIC ACTIVITY: DEGRADATION OF BLOOD COAGULATION FACTORS VA
CC AND VIIIA.
CC - SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED
CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE
CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A
CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN. THIS
CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS
CC STRONGLY PROMOTED BY THROMBOMODULIN.
CC - TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.
CC - PTM: THE VITAMIN K-DEPENDENT, ENZYMAIC CARBOXYLATION OF SOME
CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.
CC - MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING
CC SITE IS NECESSARY FOR THE RECOGNITION OF THE
CC THROMBIN-THROMBOMODULIN COMPLEX.
CC - SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02435; AAA30685.1; -
DR PIR: A00928; KXBO.
DR HSSP: P04070; 1AOT.
DR MEROPS: S01.218; -
DR InterPro: IPR000152; -
DR InterPro: IPR000294; -
DR InterPro: IPR000561; -
DR InterPro: IPR001254; -
DR InterPro: IPR001881; -
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00089; trypsin; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Blood coagulation; Glycoprotein; Plasma; Serine protease; Liver;
KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;
KW EGF-like domain; Repeat; Endothelial cell; Hydrolyase; Signal.
FT NON_TER 1 1
FT SIGNAL <1 29
FT PROPEP 30 39
FT CHAIN 40 194
FT PEPTIDE 197 456
FT DOMAIN 197 210
FT DOMAIN 94 129
FT DOMAIN 133 173
FT DOMAIN 211 456
FT MOD_RES 45 45
FT MOD_RES 46 46
FT MOD_RES 53 53
FT MOD_RES 55 55
FT MOD_RES 58 58
FT MOD_RES 59 59
FT MOD_RES 62 62
FT MOD_RES 64 64
FT MOD_RES 65 65
FT MOD_RES 68 68
FT MOD_RES 74 74
FT ACT_SITE 110 110
FT ACT_SITE 252 252
FT ACT_SITE 298 298
FT ACT_SITE 397 397
FT DISULFID 56 61
FT DISULFID 89 108
FT DISULFID 98 103
FT DISULFID 102 117
FT DISULFID 119 128
FT DISULFID 137 148
FT DISULFID 144 157
FT DISULFID 159 172
FT DISULFID 180 318
FT DISULFID 237 253
FT DISULFID 368 382
FT DISULFID 393 421
FT CARBOHYD 136 136
FT CARBOHYD 289 289
FT CARBOHYD 350 350
FT CARBOHYD 366 366
FT VARIANT 82 82
FT CONFLICT 455 456
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
F -> K.
VP -> PV (IN REF. 4).

ID FA10_BOVIN STANDARD; PRT; 492 AA.
AC P00743;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE COAGULATION FACTOR X PRECURSOR (EC 3.4.21.6) (STUART FACTOR).
GN F10.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 1-487 FROM N.A.
RX MEDLINE=84247315; PubMed=6330671;
RA Fung M.R., Campbell R.M., McGillivray R.T.A.;
RT "Blood coagulation factor X mRNA encodes a single polypeptide chain
RT containing a prepro leader sequence.";
RL Nucleic Acids Res. 12:4481-4492(1984).
RN [2]
RP SEQUENCE OF 41-180.
RX MEDLINE=80130563; PubMed=6766735;
RA Enfield D.L., Ericsson L.H., Fujikawa K., Walsh K.A., Neurath H.,
RA "Titan K.":
RT "Amino acid sequence of the light chain of bovine factor XI (Stuart
RT factor).";
RL Biochemistry 19:659-667(1980).
RN [3]
RP REVISION TO 103.
RX MEDLINE=83308813; PubMed=6688526;
RA McMullen B.A., Fujikawa K., Kistiel W.;
RT "The occurrence of beta-hydroxyaspartic acid in the vitamin
RT K-dependent blood coagulation zymogens.";
RL Biochem. Biophys. Res. Commun. 115:8-14(1983).
RN [4]
RP SEQUENCE OF 183-492, CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RX MEDLINE=7603065; PubMed=1059093;
RA "Titan K., Fujikawa K., Enfield D.L., Ericsson L.H., Walsh K.A.,
RA Neurath H.":
RT "Bovine factor XI (Stuart factor): amino-acid sequence of heavy
RT chain.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:3082-3086(1975).
RN [5]
RP SEQUENCE OF 183-233, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=94062825; PubMed=8243461;
RA Inoue K., Morita T.;
RT "Identification of O-linked oligosaccharide chains in the activation
RT peptides of blood coagulation factor X. The role of the carbohydrate
RT moieties in the activation of factor X.";
RL Eur. J. Biochem. 218:153-163(1993).
RN [6]
RP ACTIVE SITE.
RX MEDLINE=73053314; PubMed=4264286;
RA "Titan K., Hermanson M.A., Fujikawa K., Ericsson L.H., Walsh K.A.,
RA Neurath H., Davie E.W.":
RT "Bovine factor X 1a (activated Stuart factor). Evidence of homology
RT with mammalian serine proteases.";
RL Biochemistry 11:4899-4903(1972).
RN [7]
RP PROCESSING.
RX MEDLINE=76053121; PubMed=1059122;
RA Fujikawa K., Titan K., Davie E.W.;
RT "Activation of bovine factor X (Stuart factor): conversion of factor
RT Xa-alpha to factor Xa-beta.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:3359-3363(1975).
RN [8]
RP CALCIUM-BINDING DATA.
RX MEDLINE=84185716; PubMed=6546930;
RA Sugo T., Björck I., Holmgren A., Stenflo J.;
RT "Calcium-binding properties of bovine factor X lacking the gamma-
RT carboxyglutamic acid-containing region.";
RL J. Biol. Chem. 259:5705-5710(1984).
RN [9]

RP SULFATE-BINDING.
RX MEDLINE=86140210; PubMed=3949800;
RA Morita T., Jackson C.M.;
RT "Localization of the structural difference between bovine blood
RT coagulation factors XI and X2 to tyrosine 18 in the activation
RT peptide.";
RL J. Biol. Chem. 261:4008-4014(1986).
RN [10]
RP STRUCTURE BY NMR OF 85-126.
RX MEDLINE=91084483; PubMed=261466;
RA Selander M., Persson E., Stenflo J., Drakenberg T.;
RT "1H NMR assignment and secondary structure of the Ca2(+)-free form of
RT the amino-terminal epidermal growth factor like domain in coagulation
RT factor X.";
RL Biochemistry 29:8111-8118(1990).
RN [11]
RP STRUCTURE BY NMR OF 85-126.
RX MEDLINE=92329412; PubMed=1627540;
RA Ullner M., Selander M., Persson E., Stenflo J., Drakenberg T.,
RA Teلمان O.;
RT "Three-dimensional structure of the apo form of the N-terminal
RT EGF-like module of blood coagulation factor X as determined by NMR
RT spectroscopy and simulated folding.";
RL Biochemistry 31:5974-5983(1992).
RN [12]
RP STRUCTURE BY NMR OF 85-126.
RX MEDLINE=92406922; PubMed=1527084;
RA Selander M., Drakenberg T., Ullner M., Persson E., Teلمان O.,
RA Stenflo J., Drakenberg T.;
RT "How an epidermal growth factor (EGF)-like domain binds calcium. High
RT resolution NMR structure of the calcium form of the NH2-terminal EGF-
RT like domain in coagulation factor X.";
RL J. Biol. Chem. 267:19642-19649(1992).
RN [13]
RP STRUCTURE BY NMR OF 41-126.
RX MEDLINE=96387194; PubMed=8794734;
RA Sunnerhagen M., Olaf G.A., Stenflo J., Forsen S., Drakenberg T.,
RA Tremella J.;
RT "The relative orientation of Gla and EGF domains in coagulation
RT factor X is altered by Ca2+ binding to the first EGF domain. A
RT combined NMR-small angle X-ray scattering study.";
RL Biochemistry 35:11547-11559(1996).
RN [14]
RP FUNCTION: FACTOR XA IS A VITAMIN K-DEPENDENT GLYCOPROTEIN THAT
CC CONVERTS PROTHROMBIN TO THROMBIN IN THE PRESENCE OF FACTOR VA,
CC CA2+, AND PHOSPHOLIPID DURING BLOOD CLOTTING.
CC CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-1-THR AND THEN
CC ARG-1-THE BONDS IN PROTHROMBIN TO FORM THROMBIN.
CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR
CC BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR
CC MORE DISULFIDE BONDS.
CC -1- PPM: THE VITAMIN K-DEPENDENT, ENZYMAIC CARBOXYLATION OF SOME
CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC CALCIUM.
CC -1- PPM: N- AND O-GLYCOSYLATED.
CC -1- PPM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO
CC ANOTHER SITE, BEYOND THE GLA DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X00673; CAA25286.1; -
DR PIR: A00925; EXBO.
DR PDB: IAP0; 3I-JAN-94.

CC CA++ AND PHOSPHOLIPID DURING BLOOD CLOTTING.
CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-|-THR AND THEN
CC ARG-|-ILE BONDS IN PROTHROMBIN TO FORM THROMBIN.
CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE-CHAIN PRECURSOR
CC BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR
CC MORE DISULFIDE BONDS.
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMAIC CARBOXYLATION OF SOME
CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC CALCIUM.
CC -1- PTM: N- AND O-GLYCOSYLATED.
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE
CC INTERINSTRIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY).
CC -1- SIMILARITY: CONTRAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K03194; AAA52490.1; -
DR EMBL: M57285; AAA52421.1; -
DR EMBL: L29433; AAA52764.1; -
DR EMBL: L00390; AAA52764.1; JOINED.
DR EMBL: L00391; AAA52764.1; JOINED.
DR EMBL: L00392; AAA52764.1; JOINED.
DR EMBL: L00393; AAA52764.1; JOINED.
DR EMBL: L00394; AAA52764.1; JOINED.
DR EMBL: L00395; AAA52764.1; JOINED.
DR EMBL: L00396; AAA52764.1; JOINED.
DR EMBL: M22613; AAA51984.1; -
DR PIR: A00924; EXHU.
DR PIR: A25853; A25853.
DR PIR: A24478; A24478.
DR PDB: 1HCG; 08-MAY-95.
DR PDB: 1FAX; 29-OCT-97.
DR PDB: 1FXY; 17-JUN-98.
DR PDB: 1XKA; 23-MAR-99.
DR PDB: 1XKB; 23-MAR-99.
DR MEROPS: S01.216; -
DR GlycosultedB: P00742; -
DR Carbank; CCSD: 29393; -
DR MIM: 134530; -
DR InterPro: IPR000152; -
DR InterPro: IPR000294; -
DR InterPro: IPR000561; -
DR InterPro: IPR001254; -
DR InterPro: IPR001314; -
DR InterPro: IPR001881; -
DR InterPro: IPR002383; -
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00594; g1a; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00001; GLABLOOD.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PROSITE: PS00010; ASX-HYDROXYL; 1.
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01187; EGF_CA; 1.
KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW Signal; Zymogen; EGF-like domain; Repeat; 3D-structure.
FT SIGNAL 1 40
FT PROPEP 1 40
FT CHAIN 41 179 FACTOR X LIGHT CHAIN.

FT CHAIN 183 488
FT PROPEP 183 234
FT CHAIN 235 488
FT DOMAIN 86 122
FT DOMAIN 125 165
FT DOMAIN 235 488
FT MOD_RES 46 46
FT MOD_RES 47 47
FT MOD_RES 54 54
FT MOD_RES 55 56
FT MOD_RES 59 59
FT MOD_RES 60 60
FT MOD_RES 65 65
FT MOD_RES 66 66
FT MOD_RES 69 69
FT MOD_RES 72 72
FT MOD_RES 79 79
FT MOD_RES 103 103
FT CARBOHYD 199 199
FT CARBOHYD 211 211
FT CARBOHYD 221 221
FT CARBOHYD 231 231
FT ACT_SITE 276 276
FT ACT_SITE 322 322
FT ACT_SITE 419 419
FT ACT_SITE 90 101
FT DISULFID 95 110
FT DISULFID 112 121
FT DISULFID 129 140
FT DISULFID 136 149
FT DISULFID 151 164
FT DISULFID 172 342
FT DISULFID 241 246
FT DISULFID 261 277
FT DISULFID 390 404
FT DISULFID 415 443
FT CONFLICT 285 288
FT CONFLICT 442 442
SQ SEQUENCE 488 AA; 54731 MW; F81D5746AF4797AF CRC64;

Query Match 50.0%; Score 100; DB 1; Length 488;
Best Local Similarity 38.6%; Pred. No. 3; Le-09;
Matches 17; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 ANSEFLXLRHGSIXKXCIXXICDFXXAKXIFQVNDPLAFWSKH 44
Db 41 ANSEFLXLRHGSIXKXCIXXICDFXXAKXIFQVNDPLAFWSKH 84

RESULT 9
ID FA7_HUMAN STANDARD; PRT; 466 AA.
AC P08709;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE COAGULATION FACTOR VII PRECURSOR (BC 3.4.21.21) (SERUM PROTHROMBIN
DE CONVERSION ACCELERATOR) (EPTACOG ALFA).
GN F7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86205965; PubMed=3486420;
RA Hagen F.S., Gray C.L., O'Hara P.J., Grant F.J., Saari G.C.,
RA Woodbury R.G., Hart C.E., Insley M.Y., Kistiel W., Kutrachi K.,
RA Davie E.W.;
RT "Characterization of a cDNA coding for human factor VII.";
RN Proc. Natl. Acad. Sci. U.S.A. 83:2412-2416(1986).
[2]

RP SEQUENCE FROM N.A.
 RX MEDLINE-87260948; PubMed-3037537;
 RA O'Hara P.J., Grant F.J., Haldeman B.A., Gray C.L., Insley M.Y.,
 RA Hagen F.S., Murray M.J.;
 RT "Nucleotide sequence of the gene coding for human factor VII, a
 RT vitamin K-dependent protein participating in blood coagulation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:5158-5162(1987).
 RN [3]
 RP SEQUENCE OF 61-466, AND POST-TRANSLATIONAL MODIFICATIONS.
 RX MEDLINE-89088153; PubMed-3264725;
 RA Thim L., Bjoern S., Christensen M., Nicolaisen E.M., Lund-Hansen T.,
 RA Pedersen A.H., Hedner U.;
 RT "Amino acid sequence and posttranslational modifications of human
 RT factor VIIa from plasma and transfected baby hamster kidney cells.";
 RL Biochemistry 27:7785-7793(1988).
 RN [4]
 RP CARBOHYDRATE-LINKAGE SITES SER-112 AND SER-120.
 RX MEDLINE-91250411; PubMed-1904059;
 RA Bjoern S., Foster D.C., Thim L., Wlberg F.C., Christensen M.,
 RA Komiyama Y., Pedersen A.H., Kistiel W.;
 RT "Human plasma and recombinant factor VII. Characterization of O-
 RT glycosylations at serine residues 52 and 60 and effects of site-
 RT directed mutagenesis of serine 52 to alanine.";
 RL J. Biol. Chem. 266:11051-11057(1991).
 RN [5]
 RP STRUCTURE OF CARBOHYDRATE ON SER-112.
 RX MEDLINE-90062160; PubMed-2511201;
 RA Nishimura H., Kawabata S., Kistiel W., Hase S., Ikenaka T., Takao T.,
 RA Shimonishi Y., Iwanaga S.;
 RT "Identification of a disaccharide (xy1-Glc) and a trisaccharide
 RT (xy12-Glc) O-glycosidically linked to a serine residue in the first
 RT epidermal growth factor-like domain of human factors VII and IX and
 RT protein Z and bovine protein Z.";
 RL J. Biol. Chem. 264:20320-20325(1989).
 RN [6]
 RP STRUCTURE OF CARBOHYDRATE ON SER-112.
 RX MEDLINE-91344709; PubMed-2129367;
 RA Iwanaga S., Nishimura H., Kawabata S., Kistiel W., Hase S., Ikenaka T.;
 RT "A new trisaccharide sugar chain linked to a serine residue in the
 RT first EGF-like domain of clotting factors VII and IX and protein Z.";
 RL Adv. Exp. Med. Biol. 281:121-131(1990).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
 RX MEDLINE-96175641; PubMed-8598903;
 RA Banner D.W., D'Arcy A., Chene C., Winkler F.K., Guha A.,
 RA Konigsberg W.H., Nemerson Y., Kirchhofer D.;
 RT "The crystal structure of the complex of blood coagulation factor
 RT VIIa with soluble tissue factor.";
 RL Nature 380:41-46(1996).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF FVIIA IN COMPLEX WITH TF.
 RX MEDLINE-99126538; PubMed-9925787;
 RA Zhang E., St. Charles R., Tulinsky A.;
 RT "Structure of extracellular tissue factor complexed with factor VIIa
 RT inhibited with a BPTI mutant.";
 RL J. Mol. Biol. 285:2089-2104(1999).
 RN [9]
 RP STRUCTURE BY NMR OF 105-145.
 RX MEDLINE-98367502; PubMed-9692950;
 RA Mukanen A., Finn B.E., Gilpert G.P., Forsen S., Stenflo J.,
 RA Drekenberg T.;
 RT "Solution structure of the N-terminal EGF-like domain from human
 RT factor VII.";
 RL Biochemistry 37:10605-10615(1998).
 RN [10]
 RP VARIANT GLN-364.
 RX MEDLINE-91300046; PubMed-2070047;
 RA O'Brien D.P., Gale K.M., Anderson J.S., McVey J.H., Miller G.J.,
 RA Meade T.W., Tuddenham E.G.D.;
 RT "Purification and characterization of factor VII 304-Gln: a variant
 RT molecule with reduced activity isolated from a clinically unaffected
 RT male.";
 RL Blood 78:132-140(1991).
 RN [11]
 RP VARIANTS GLN-364 AND PHE-370.
 RX MEDLINE-92340074; PubMed-1634227;
 RA Marchetti G., Patraccchini P., Gemmati D., Derosa V., Pinotti M.,
 RA Rodorigo G., Casanato A., Girolami A., Bernardi F.;
 RT "Detection of two missense mutations and characterization of a repeat
 RT polymorphism in the factor VII gene (F7).";
 RL Hum. Genet. 89:497-502(1992).
 RN [12]
 RP VARIANT TYR-238.
 RX MEDLINE-93378811; PubMed-8364544;
 RA Marchetti G., Ferrari M., Patraccchini P., Redaelli R., Bernardi F.;
 RT "A missense mutation (178Cys->Tyr) and two neutral dimorphisms
 RT (115His and 333Ser) in the human coagulation factor VII gene.";
 RL Hum. Mol. Genet. 2:1055-1056(1993).
 RN [13]
 RP VARIANTS.
 RX MEDLINE-94061028; PubMed-8242057;
 RA Takamiya O., Kempali-Cook G., Marin D.M.A., Cooper D.N.,
 RA von Felten A., Mell E., Hahn I., Prangnell D.R., Lunley H.,
 RA Tuddenham E.G.D., McVey J.H.;
 RT "Detection of missense mutations by single-strand conformational
 RT polymorphism (SSCP) analysis in five dysfunctional variants of
 RT coagulation factor VII.";
 RL Hum. Mol. Genet. 2:1355-1359(1993).
 RN [14]
 RP VARIANTS CHARLOTTE GLN-139 AND GLN-212.
 RX MEDLINE-94264305; PubMed-8204879;
 RA Ching S., Clarke B., Sridhara S., Chu K., Friedman P., Vandusen W.,
 RA Roberts H.R., Blajchman M., Monroe D.M., High K.A.;
 RT "Severe factor VII deficiency caused by mutations abolishing the
 RT cleavage site for activation and altering binding to tissue factor.";
 RL Blood 83:3524-3535(1994).
 RN [15]
 RP VARIANT VAL-354.
 RX MEDLINE-95072589; PubMed-7981691;
 RA Bernardi F., Castaman G., Redaelli R., Pinotti M., Lunghi B.,
 RA Rodeghiero F., Marchetti G.;
 RT "Topologically equivalent mutations causing dysfunctional coagulation
 RT factors VII (294Ala->Val) and X (334Ser->Pro).";
 RL Hum. Mol. Genet. 3:1175-1177(1994).
 RN [16]
 RP VARIANT MET HIS-307.
 RX MEDLINE-95064662; PubMed-7974346;
 RA Ohlwa M., Hayashi T., Mada H., Minamikawa K., Shitake S.,
 RA Suzuki K.;
 RT "Factor VII Met: homozygous asymptomatic type I deficiency caused by
 RT an amino acid substitution of His (CAC) for Arg(247) (CGC) in the
 RT catalytic domain.";
 RL Thromb. Haemost. 71:773-777(1994).
 RN [17]
 RP VARIANT MET-419.
 RX MEDLINE-96247510; PubMed-8652821;
 RA Arbini A., Mannucci P.M., Bauer K.A.;
 RT "A Thr359Met mutation in factor VII of a patient with a hereditary
 RT deficiency causes defective secretion of the molecule.";
 RL Blood 87:5085-5094(1996).
 RN [18]
 RP VARIANTS W-283; K-325; V-358; Q-364; E-402 AND Q-413.
 RX MEDLINE-97001216; PubMed-8844208;
 RA Bernardi F., Castaman G., Pinotti M., Ferraresi P., di Iasio M.G.,
 RA Lunghi B., Rodeghiero F., Marchetti G.;
 RT "Mutation pattern in clinically asymptomatic coagulation factor VII
 RT deficiency.";
 RL Hum. Mutat. 8:108-115(1996).
 RN [19]
 RP VARIANT VAL-304.
 RX MEDLINE-97037613; PubMed-8883260;
 RA Tanary H., Fromovich Y., Salmon L., Reich Z., Dym O., Lanir N.,
 RA Brenner B., Paz M., Luder A.S., Blau O., Korostishevsky M.,
 RA Zaitov R., Seligson U.;
 RT "Ala244Val is a common, probably ancient mutation causing factor VII
 RT deficiency in Moroccan and Iranian Jews.";

```
RL Thromb. Haemost. 76:283-291(1996).
RN [20]
RP VARIANTS MALTA THR-194 AND VAL-304.
RX MEDLINE=98112461; PubMed=9452082;
RA Ashnawi C., Scerif C., Galdies R., Aguilina A., Felice A.E.;
RT "Two new missense mutations (P134T and A244V) in the coagulation
   factor VII gene."
RL Hum. Mutat. Suppl. 1:S189-S191(1998).
CC -1- FUNCTION: CIRCULATES IN THE BLOOD IN A ZYMOGEN FORM. FACTOR VII IS
   CONVERTED TO FACTOR VIIA BY FACTOR XA, FACTOR XIa, FACTOR IXa, OR
   THROMBIN BY MINOR PROTEOLYSES. IN THE PRESENCE OF TISSUE FACTOR
   AND CALCIUM IONS, FACTOR VIIa THEN CONVERTS FACTOR X TO FACTOR Xa
   BY LIMITED PROTEOLYSIS. FACTOR VIIa WILL ALSO CONVERT FACTOR IX TO
   FACTOR IXa IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM.
CC -1- CATALYTIC ACTIVITY: HYDROLYSES ONE ARG-1-ILE BOND IN FACTOR X TO
   FORM FACTOR Xa.
CC -1- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED
   BY A DISULFIDE BOND.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
   GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
   CALCIUM.
CC -1- DISEASE: DEFECTS IN F7 CAN CAUSE COAGULOPATHY.
CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES NISTASE OR NOVOSEVEN
   (NOVO NORDISK). USED FOR THE TREATMENT OF BLEEDING EPISODES IN
   HEMOPHILIA A OR B PATIENTS WITH ANTIBODIES TO COAGULATION FACTORS
   VIII OR IX.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
   TRYPSIN FAMILY.
CC -1- DATABASE: NAME-Factor VII mutation database;
   WWW=http://europium.mrc.rpms.ac.uk/ust/WWW/webPages/FVII/database.dir/tilepag
   -----
Query Match          46.5%; Score 93; DB 1; Length 466;
Best Local Similarity 46.3%; Pred. No. 4.5e-08;
Matches 19; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 ANSEFLXMRHSLRXCIXICDFEXAKXIFQNDVTLAFW 41
   ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 ANAFLELRPGSLRECKEKGCSFEAREIKDAERKFLFW 101

RESULT 10
FA10_RABIT STANDARD; PRT; 490 AA.
AC 019045;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE COAGULATION FACTOR X PRECURSOR (EC 3.4.21.6) (STUART FACTOR).
GN F10.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97256311; PubMed=9101642;
RA Pendurthi U.R., Anderson K.D., James H.L.;
RT "Characterization of a full-length cDNA for rabbit factor X."
RL Thromb. Res. 85:503-514(1997).
CC -1- FUNCTION: FACTOR Xa IS A VITAMIN K-DEPENDENT GLYCOPROTEIN THAT
   CONVERTS PROTHROMBIN TO THROMBIN IN THE PRESENCE OF FACTOR Va,
   Ca++ AND PHOSPHOLIPID DURING BLOOD CLOTTING.
CC -1- CATALYTIC ACTIVITY: PREFERENTIALLY CLEAVAGE: ARG-1-THR AND THEN
   ARG-1-ILE BONDS IN PROTHROMBIN TO FORM THROMBIN.
CC -1- SUBUNIT: THE TWO CHAINS ARE FORMED FROM A SINGLE CHAIN PRECURSOR
   BY THE EXCISION OF TWO ARG RESIDUES AND ARE HELD TOGETHER BY 1 OR
   MORE DISULFIDE BONDS.
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMATIC CARBOXYLATION OF SOME
   GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
   CALCIUM (BY SIMILARITY).
```

```

FT  MOD_RES      103      103      SIMILARITY).
FT  ACT_SITE     274      274      HYDROXYLATION (BY SIMILARITY).
FT  ACT_SITE     320      320      CHARGE RELAY SYSTEM.
FT  ACT_SITE     417      417      CHARGE RELAY SYSTEM.
FT  DISULFID     90      101      BY SIMILARITY.
FT  DISULFID     95      110      BY SIMILARITY.
FT  DISULFID     112      121      BY SIMILARITY.
FT  DISULFID     129      140      BY SIMILARITY.
FT  DISULFID     136      149      BY SIMILARITY.
FT  DISULFID     151      164      BY SIMILARITY.
FT  DISULFID     172      340      INTERCHAIN (BY SIMILARITY).
FT  DISULFID     239      244      BY SIMILARITY.
FT  DISULFID     259      275      BY SIMILARITY.
FT  DISULFID     388      402      BY SIMILARITY.
FT  DISULFID     413      441      BY SIMILARITY.
FT  CARBOHYD     61      61      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD     187      187      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD     205      205      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE      490 AA; 53965 MW; 3A39FA85AF2A6D11 CRC64;

Query Match      46.58; Score 93; DB 1; Length 490;
Best Local Similarity 38.6%; Pred. No. 4,76-08;
Matches 17; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Oy  1  ANSFLXLRHGSIXKXICDFFXAKKXICPNVDTLAFMKH 44
Db  41  ANSFLKLGKLRKMEKNCSEAELEVEDREKTEFMNKKY 84

RESULT  11
ID  FA7_BOVIN      STANDARD:      PRT:  407 AA.
AC  P22457;
DT  01-AUG-1991 (Rel. 19, Created)
DT  01-AUG-1991 (Rel. 19, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  COAGULATION FACTOR VII (EC 3.4.21.21) (SERUM PROTHROMBIN CONVERSION
DE  ACCELERATOR).
CN  F7.
OS  Bos taurus (Bovine).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovidae; Bovinae; Bos.
OX  NCBI_Taxid=9913;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=89008362; PubMed=3049594;
RA  Takeya H., Kawabata S., Nakagawa K., Yamamichi Y., Miyata T.,
RA  Iwanaga S.;
RT  "Bovine factor VII. Its purification and complete amino acid
RT  sequence."
RL  J. Biol. Chem. 263:14868-14877(1988).
RN  [2]
RP  STRUCTURE OF CARBOHYDRATE ON SER-52.
RX  MEDLINE=89213999; PubMed=3149637;
RA  Hase S., Kawabata S., Nishimura H., Takeya H., Sueyoshi T.,
RA  Miyata T., Iwanaga S., Takao T., Shimonishi Y., Ikenaka T.;
RT  "A new trisaccharide sugar chain linked to a serine residue in bovine
RT  blood coagulation factors VII and IX."
RL  J. Biochem. 104:867-868(1988).
RN  [3]
RP  STRUCTURE OF CARBOHYDRATE ON SER-52.
RX  MEDLINE=91344709; PubMed=2129367;
RA  Iwanaga S., Nishimura H., Kawabata S., Kistiel W., Hase S., Ikenaka T.;
RT  "A new trisaccharide sugar chain linked to a serine residue in the
RT  first EGF-like domain of clotting factors VII and IX and protein Z."
RL  Adv. Exp. Med. Biol. 281:121-131(1990).
CC  -1- FUNCTION: CIRCULATES IN THE BLOOD IN A ZYMOGEN FORM. FACTOR VII IS
CC  CONVERTED TO FACTOR VIIA BY FACTOR XA, FACTOR XIa, FACTOR IXa, OR
CC  THROMBIN BY MINOR PROTEOLYSIS. IN THE PRESENCE OF TISSUE FACTOR
CC  AND CALCIUM IONS, FACTOR VIIa THEN CONVERTS FACTOR X TO FACTOR Xa

```

```

CC  BY LIMITED PROTEOLYSIS. FACTOR VIIa WILL ALSO CONVERT FACTOR IX TO
CC  FACTOR IXa IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM.
CC  -1- CATALYTIC ACTIVITY: HYDROLYSES ONE ARG-1-ILE BOND IN FACTOR X TO
CC  FORM FACTOR Xa.
CC  -1- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED
CC  BY A DISULFIDE BOND.
CC  -1- TISSUE SPECIFICITY: PLASMA.
CC  -1- PTM: THE VITAMIN K-DEPENDENT. ENZYMAIC CARBOXYLATION OF SOME
CC  GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND
CC  CALCIUM.
CC  -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC  TRYPSIN FAMILY.
DR  PIR: A31979; A31979.
DR  HSRP: P00740; 1IXA.
DR  MEROPS: S01.215. -.
DR  InterPro: IPR000152; -.
DR  InterPro: IPR000294; -.
DR  InterPro: IPR000561; -.
DR  InterPro: IPR001254; -.
DR  InterPro: IPR001314; -.
DR  InterPro: IPR001438; -.
DR  InterPro: IPR001881; -.
DR  InterPro: IPR002383; -.
DR  Pfam: PF00008; EGF_2.
DR  Pfam: PF00594; gla_1.
DR  Pfam: PF00089; trypsin_1.
DR  PRINTS: PRO0001; GLABLOOD.
DR  PRINTS: PRO0010; EGFIBLOOD.
DR  PRINTS: PRO0722; CHYMOTRYPSIN.
DR  PROSITE: PS00010; ASX_HYDROXYL. 1.
DR  PROSITE: PS00011; GLU_CARBOXYLATION; 1.
DR  PROSITE: PS00022; EGF_1; 1.
DR  PROSITE: PS00134; TRYPSIN_HIS_1.
DR  PROSITE: PS00135; TRYPSIN_SER; 1.
DR  PROSITE: PS01186; EGF_2; 2.
DR  PROSITE: PS01187; EGF_CA; 1.
KW  Hydrolyase; Serine protease; Blood coagulation; zymogen; Glycoprotein;
KW  Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;
KW  EGF-like domain; Repeat.
FT  CHAIN      1      152      FACTOR VII LIGHT CHAIN.
FT  CHAIN      153      407      FACTOR VII HEAVY CHAIN.
FT  DOMAIN      6      35      EGF-RICH.
FT  DOMAIN      46      82      EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT  DOMAIN      87      128      EGF-LIKE 2.
FT  DOMAIN      153      407      CATALYTIC.
FT  SITE      152      153      CLEAVAGE (BY FACTOR Xa, FACTOR XIa,
FT  FACTOR IXa, OR THROMBIN).
FT  ACT_SITE     193      193      BY SIMILARITY.
FT  ACT_SITE     242      242      BY SIMILARITY.
FT  ACT_SITE     344      344      BY SIMILARITY.
FT  BINDING      338      338      SUBSTRATE (BY SIMILARITY).
FT  DISULFID     17      22      BY SIMILARITY.
FT  DISULFID     50      61      BY SIMILARITY.
FT  DISULFID     55      70      BY SIMILARITY.
FT  DISULFID     72      81      BY SIMILARITY.
FT  DISULFID     91      102      BY SIMILARITY.
FT  DISULFID     98      112      BY SIMILARITY.
FT  DISULFID     114      127      BY SIMILARITY.
FT  DISULFID     135      162      BY SIMILARITY.
FT  DISULFID     159      264      BY SIMILARITY.
FT  DISULFID     178      194      BY SIMILARITY.
FT  DISULFID     310      329      BY SIMILARITY.
FT  DISULFID     340      368      BY SIMILARITY.
FT  MOD_RES      7      7      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      6      6      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      14      14      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      16      16      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      19      19      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      20      20      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      25      25      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      26      26      GAMMA-CARBOXYGLUTAMIC ACID.
FT  MOD_RES      29      29      GAMMA-CARBOXYGLUTAMIC ACID.

```


DB 44 ANSGLEELRGNTERECVEBOCSYEAEFALESPODITVEMAKY 88

RESULT 14

PRTS_MACMU STANDARD: PRT: 649 AA.

ID PRTS_MACMU STANDARD: PRT: 649 AA.

AC Q28520; 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE VITAMIN K-DEPENDENT PROTEIN S PRECURSOR (FRAGMENT).

GN PROSI OR PROS.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.

OX NCBI_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA MEDLINE=95134217; PubMed=7832752;

RA Greenard J.S., Fernandez J.A., Radtke K.P., Griffin J.H.;

RT "Identification of candidate residues for interaction of protein S with C4b binding protein and activated protein C.";

RL Biochem. J. 305:397-403(1995).

CC -1- FUNCTION: ANTICOAGULANT PLASMA PROTEIN; IT IS A COFACTOR TO ACTIVATED PROTEIN C IN THE DEGRADATION OF COAGULATION FACTORS VA AND VIIIA. IT HELP TO PREVENT COAGULATION AND STIMULATING FIBRINOLYSIS.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.

CC CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: L31380; AAA70376.1; ALT_INT.

DR HSP: P00740; IGFH.

DR InterPro: IPR000152; -

DR InterPro: IPR000294; -

DR InterPro: IPR000561; -

DR InterPro: IPR001791; -

DR InterPro: IPR001881; -

DR Pfam: PF00008; EGF_4.

DR Pfam: PF00594; gla; 1.

DR Pfam: PF00594; gla; 1.

DR Pfam: PF00594; gla; 1.

DR PROSITE: PS00010; ASX_HYDROXYL; 4.

DR PROSITE: PS00011; GLU CARBOXYLATION; 1.

DR PROSITE: PS00022; EGF_1; 1.

DR PROSITE: PS01186; EGF_2; 3.

DR PROSITE: PS01187; EGF_CA; 3.

KW Plasma; Gamma-carboxyglutamic acid; Calcium; Vitamin K; Zymogen;

KW Blood coagulation; Hydroxylation; Glycoprotein; Repeat;

KW EGF-like domain.

FT NON_TER 1 1

FT PROPEP <1 14

FT CHAIN 15 649

FT DOMAIN 15 59

FT DOMAIN 60 89

FT DOMAIN 90 128

FT DOMAIN 130 173

FT DOMAIN 174 215

FT DOMAIN 216 236

FT MOD_RES 20 20

FT MOD_RES 21 21

FT MOD_RES 21 21

BY SIMILARITY.

VITAMIN K-DEPENDENT PROTEIN S.

GLA.

THROMBIN-SENSITIVE.

EGF-LIKE 1.

EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).

EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).

EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).

GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

GAMMA-CARBOXYGLUTAMIC ACID (BY SIMILARITY).

FT MOD_RES 28 28

FT MOD_RES 30 30

FT MOD_RES 33 33

FT MOD_RES 34 34

FT MOD_RES 39 39

FT MOD_RES 40 40

FT MOD_RES 43 43

FT MOD_RES 46 46

FT MOD_RES 50 50

FT MOD_RES 109 109

FT DISULFID 94 107

FT DISULFID 99 116

FT DISULFID 118 127

FT DISULFID 134 148

FT DISULFID 144 157

FT DISULFID 159 172

FT DISULFID 178 190

FT DISULFID 185 199

FT DISULFID 201 214

FT DISULFID 220 229

FT DISULFID 225 238

FT DISULFID 240 255

FT CARBOHYD 472 482

FT CARBOHYD 482 482

FT CARBOHYD 503 503

SO SEQUENCE 649 AA; 72402 MW; 5C7C13D31C079B6B CRC64;

Query Match 41.0%; Score 82; DB 1; Length 649;

Best Local Similarity 36.4%; Pred. No. 4.7e-06;

Matches 16; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGLRXCIXICDFXAKXIFQNDPDLAFMSKH 44

DB 15 ANSMLEETKQGNLECEIEICNKEAREVFNDEPDTYFPKY 58

RESULT 15

PRTS_HUMAN STANDARD: PRT: 676 AA.

ID PRTS_HUMAN STANDARD: PRT: 676 AA.

AC P07225; Q15518; 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE VITAMIN K-DEPENDENT PROTEIN S PRECURSOR.

GN PROSI OR PROS.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91084444; PubMed=2148110;

RA Schmidel D.R., Tatro A.V., Phelps L.G., Tomczak J.A., Long G.L.;

RT "Organization of the human protein S genes.";

RT Biochemistry 29:7845-7852(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Ploos van Amstel H.K.;

RL Submitted (AUG-1987) to the EMBL/GenBank/DBJ databases.

RN [3]

RX MEDLINE=91084445; PubMed=2148111;

FT	DISULFID	161	175	BY SIMILARITY.
FT	DISULFID	171	184	BY SIMILARITY.
FT	DISULFID	186	199	BY SIMILARITY.
FT	DISULFID	205	217	BY SIMILARITY.
FT	DISULFID	212	226	BY SIMILARITY.
FT	DISULFID	228	241	BY SIMILARITY.
FT	DISULFID	247	256	BY SIMILARITY.
FT	DISULFID	252	265	BY SIMILARITY.
FT	DISULFID	267	282	BY SIMILARITY.
FT	CARBOHYD	499	499	N-LINKED (GLCNAC. . .); EXCEPT IN VARIANT HEERLEN.
FT	CARBOHYD	509	509	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	530	530	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	196	196	K -> E (IN TOKUSHIMA; PROS DEFICIENCY).
FT	VARIANT	258	258	/FTId-VAR_005566. N -> S (IN PROS DEFICIENCY).
FT	VARIANT	501	501	/FTId-VAR_005567 S -> P (IN HEERLEN).
FT	CONFLICT	11	11	/FTId-VAR_005568. L -> P (IN REF. 5).
FT	CONFLICT	26	26	F -> L (IN REF. 5).
SO	SEQUENCE	676 AA;	75122 MW;	2B88A04F85403F25 CRC64;

Query Match 41.08; Score 82; DB 1; Length 676;
 Best Local Similarity 36.4%; Pred. No. 4.9e-06;
 Matches 16; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

OY 1 ANSFLXXLRHGSIXXICDFFXAKXIFQONVDDTLAFWSKH 44
 DB 42 ANSLLEETKQGNLECEIEECNKEAREVFNDEPTDYFPKY 85

Search completed: November 8, 2001, 08:56:10
 Job time: 174 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2001, 08:50:01 ; Search time 13.45 Seconds
(without alignments)
249.196 Million cell updates/sec

Title: US-09-497-591-1

Perfect score: 200
Sequence: 1 ANSFLXXLRHSGSLXRCIXX.....XXAKXLFQNVDTLAFMSKH 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	88.0	461	1 KXHU	protein C (activat
2	147	73.5	461	1 UX0210	protein C (activat
3	146	73.0	461	1 S18994	protein C (activat
4	129	64.5	456	1 KXBO	protein C (activat
5	105	52.5	482	1 EXRT	coagulation factor
6	104	52.0	443	2 I46932	coagulation factor
7	101	50.5	492	1 EXBO	coagulation factor
8	100	50.0	488	1 EXHU	coagulation factor
9	93	46.5	466	1 KFH07	coagulation factor
10	90	45.0	407	1 KFB07	coagulation factor
11	82.5	41.2	617	2 S10511	thrombin (EC 3.4.2
12	82.5	41.2	618	2 A35827	thrombin (EC 3.4.2
13	82	41.0	642	2 S53434	plasma protein S p
14	82	41.0	676	1 KXHUS	plasma protein S p
15	81	40.5	452	1 A30351	coagulation factor
16	81	40.5	459	2 J00419	coagulation factor
17	81	40.5	466	2 S38819	plasma protein S -
18	80	40.0	622	1 TBHU	thrombin (EC 3.4.2
19	80	40.0	675	1 KXBOS	plasma protein S p
20	78	39.0	675	1 KXRTS	plasma protein S p
21	77	38.5	461	1 KFHU	coagulation factor
22	77	37.5	642	1 S53433	plasma protein S p
23	74	37.0	416	1 KFB0	coagulation factor
24	73	36.5	475	1 EXCH	coagulation factor
25	71	35.5	675	1 KXMS	plasma protein S p
26	68	34.0	625	1 TBBO	thrombin (EC 3.4.2
27	65.5	32.8	422	1 KKHU2	plasma protein 2 p
28	62	31.0	673	2 A48089	growth arrest-spec
29	61	30.5	674	2 I55476	growth potentialin

30	60	30.0	678	2 B48089	growth arrest-spec
31	59.5	29.8	396	1 KXB02	plasma protein 2 -
32	51.5	25.8	594	2 D64859	probable MAP kinase
33	51	25.5	245	2 S49045	luxP protein - Vib
34	49.5	24.8	510	2 E82918	ammonium transport
35	49.5	24.8	603	2 C96575	probable MAP kinase
36	49	24.5	422	2 T39306	mitogen-activated
37	49	24.5	730	2 G64062	primosomal replica
38	48	24.0	606	2 T40556	hypothetical prote
39	47.5	23.8	489	2 S69027	ammonium transport
40	47.5	23.8	576	2 G36763	probable MAP kinase
41	47	23.5	390	2 T14794	hypothetical prote
42	47	23.5	467	2 S36559	l2 protein - human
43	47	23.5	628	2 T04252	probable phosphati
44	47	23.5	1330	2 S49010	embryonic receptor
45	47	23.5	1333	2 I78875	receptor tyrosine

ALIGNMENTS

RESULT 1
KXHU
protein C (activated) (EC 3.4.21.69) precursor - human
N:Alternate names: autoprothrombin IIA; plasma protein C
C:Species: Homo sapiens (man)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 16-Jul-1999
C:Accession: A22331; A25426; A21781; A23789; A00927
R:Foster, D.C.; Yoshitake, S.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 4673-4677, 1985
A:Title: The nucleotide sequence of the gene for human protein C.
A:Reference number: A22331; MUID:85270390
A:Accession: A22331
A:Molecule type: DNA
A:Residues: 1-461 <FOS1>
A:Cross-references: GB:M1228; NID:g190333; PIDN:AAA60166.1; PID:g190334
R:Plutsky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.
Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986
A:Title: Evolution and organization of the human protein C gene.
A:Reference number: A25426; MUID:86120978
A:Accession: A25426
A:Molecule type: DNA
A:Residues: 1-445, 'L', 446-461 <PLU>
A:Cross-references: GB:M12712; NID:g190330; PIDN:AAA60165.1; PID:g190332
R:Foster, D.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 81, 4766-4770, 1984
A:Title: Characterization of a cDNA coding for human protein C.
A:Reference number: A21781; MUID:84272714
A:Accession: A21781
A:Molecule type: mRNA
A:Residues: 'Q', 107-461 <FOS2>
A:Cross-references: GB:K02058; NID:g190322; PIDN:AAA60164.1; PID:g190323
R:Beckmann, R.J.; Schmidt, R.J.; Santere, R.F.; Plutsky, J.; Crabtree, G.R.; Long, G.
Nucleic Acids Res. 13, 5233-5247, 1985
A:Title: The structure and evolution of a 461 amino acid human protein C precursor an
A:Reference number: A23789; MUID:85269639
A:Accession: A23789
A:Molecule type: mRNA
A:Residues: 1-461 <BEC>
A:Cross-references: GB:X02750; NID:g35689; PIDN:CMA26528.1; PID:g763120
R:Milietich, J.P.; Broze Jr., G.J.
J. Biol. Chem. 265, 11397-11404, 1990
A:Title: Beta protein C is not glycosylated at asparagine 329. The rate of translatio
A:Reference number: A44605; MUID:90293094
A:Contents: annotation; carbohydrate binding sites; activation peptide
A:Note: the alpha form of protein C is glycosylated at Asn-329, and the beta form is
R:Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A:Title: O-linked fucose is present in the first epidermal growth factor domain of fa
A:Reference number: A44606; MUID:92184750
A:Contents: annotation; beta-hydroxyaspartic acid
A:Comment: protein C is the zymogen of the vitamin K-dependent serine proteinase that
ivation of factor Va is strongly enhanced by complexing with protein S. Protein C als

C;Comment: Protein C is synthesized in the liver as a single chain precursor, which is a bin. which cleaves a dodecapeptide from the amino end of the heavy chain; this reaction, C;Genetics:

A;Gene: GDB:PROC

A;Cross-references: GDB:120317; OMIM:176860

A;Map position: 2q13-2q21

A;Intons: 24/1; 79/3; 88/1; 134/1; 179/1; 226/3; 266/1

C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology

C;Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding

F:1-32/Domain: signal sequence #status predicted <SIG>

F:27-86/Domain: Gla domain homology <Gla>

F:33-42/Domain: propeptide #status predicted <PRO>

F:43-197/Product: protein C light chain #status predicted <LCH>

F:92-131/Domain: EGF homology <EG1>

F:140-175/Domain: EGF homology <EG2>

F:200-211/Domain: activation peptide #status experimental <APT>

F:212-445/Domain: trypsin homology <TRY>

F:48,49,56,58,61,62,67,68,71/Modified site: gamma-carboxyglutamic acid (Glu) #status exp

F:59-64,92-105,101-120,122-131,140-151,147-160,162-175,183-319,238-254,373-387,398-426/D

F:106-111/Disulfide bonds: #status predicted

F:110/Binding site: carbonydrate (Thr) (covalent) #status absent

F:113/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental

F:139,290,355/Binding site: carbonydrate (Asn) (covalent) #status experimental

F:211,212/Cleavage site: Arg-Leu (thrombin) #status experimental

F:233,499,402/Active site: His, Asp, Ser #status predicted

F:371/Binding site: carbonydrate (Asn) (covalent) (partial) #status atypical

Query Match 88.0%; Score 176; DB 1; Length 461;

Best Local Similarity 77.3%; Pred. No. 5e-21;

Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGSILKRCIXICDFXKAKXIFQNVDDTLAFWSKH 44

Db 43 ANSFLERHRSLSRECEIECDPEEAKEIFQNVDDTLAFWSKH 86

RESULT 2

protein C (activated) (EC 3.4.21.69) precursor - mouse

N;Alternate names: vitamin K-dependent serine proteinase

C;Species: Mus musculus (house mouse)

C;Date: 10-Sep-1999 #sequence,revision 10-Sep-1999 #text_change 16-Jun-2000

C;Accession: JX0210

R;Tada, N.; Sato, M.; Tsujimura, A.; Iwase, R.; Hashimoto-Gotoh, T.

J. Biochem. 111, 491-495, 1992

A;Title: Isolation and characterization of a mouse protein C cDNA.

A;Reference number: JX0210; MUID:92316897

A;Accession: JX0210

A;Molecule type: mRNA

A;Residues: 1-461 <TAD>

A;Cross-references: GB:D10445; NID:g220385; PIDN:BAA01235.1; PID:g220386

A;Experimental source: liver

C;Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re

S;

C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology

C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutam

F:1-33/Domain: signal sequence #status predicted <SIG>

F:27-85/Domain: Gla domain homology <Gla>

F:34-41/Domain: propeptide #status predicted <PRO>

F:42-196,199-461/Product: protein C #status predicted <PRC>

F:42-196/Domain: light chain #status predicted <PCL>

F:91-130/Domain: EGF homology <EG1>

F:139-174/Domain: EGF homology <EG2>

F:199-461/Domain: heavy chain #status predicted <PCH>

F:199-211/Domain: activation peptide #status predicted <ACT>

F:212-461/Product: vitamin K-dependent serine proteinase #status predicted <VIT>

F:212-445/Domain: trypsin homology <TRY>

F:47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status

F:121/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

F:131-330,139-150,146-159,161-174,183-319,238-254,373-387,398-426/Disulfide bonds: #stat

F:214,290,355/Binding site: carbonydrate (Asn) (covalent) #status predicted

F:253,299,402/Active site: His, Asp, Ser #status predicted

Query Match 73.5%; Score 147; DB 1; Length 461;

Best Local Similarity 63.6%; Pred. No. 3e-16;

Matches 28; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGSILKRCIXICDFXKAKXIFQNVDDTLAFWSKH 44

Db 42 ANSFLERHRSLSRECEIECDPEEAKEIFQNVDDTLAFWSKH 85

RESULT 3

protein C (activated) (EC 3.4.21.69) precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 10-Sep-1999 #sequence,revision 10-Sep-1999 #text_change 29-Oct-1999

C;Accession: S18994; S24312

R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.

submitted to the EMBL Data Library, February 1992

A;Description: The cDNA cloning and mRNA expression of rat protein C.

A;Reference number: S18994

A;Accession: S18994

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-461 <OKA>

A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963

R;Okafuji, T.; Maekawa, K.; Nawa, K.; Marumoto, Y.

Biochim. Biophys. Acta 1131, 329-332, 1992

A;Title: The cDNA cloning and mRNA expression of rat protein C.

A;Reference number: S24312; MUID:92329550

A;Accession: S24312

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-461 <OKA2>

A;Cross-references: EMBL:X64336; NID:g56962; PIDN:CAA45617.1; PID:g56963

C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol

C;Keywords: beta-hydroxyaspartic acid; glycoprotein; hydrolase; serine proteinase

F:1-32/Domain: signal sequence #status predicted <SIG>

F:27-85/Domain: Gla domain homology <Gla>

F:33-42/Domain: propeptide #status predicted <PRO>

F:43-461/Product: protein C #status predicted <PRC>

F:91-130/Domain: EGF homology <EG1>

F:139-174/Domain: EGF homology <EG2>

F:213-445/Domain: trypsin homology <TRY>

F:47,48,55,57,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #stat

F:112/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

F:121-130,139-150,146-159,161-174,183-320,239-255,373-387,398-426/Disulfide bonds: #s

F:215,291,355/Binding site: carbonydrate (Asn) (covalent) #status predicted

F:254,300,402/Active site: His, Asp, Ser #status predicted

Query Match 73.0%; Score 146; DB 1; Length 461;

Best Local Similarity 63.6%; Pred. No. 4.3e-16;

Matches 29; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGSILKRCIXICDFXKAKXIFQNVDDTLAFWSKH 44

Db 42 ANSFLERHRSLSRECEIECDPEEAKEIFQNVDDTLAFWSKH 85

RESULT 4

KXBO

protein C (activated) (EC 3.4.21.69) precursor - bovine (fragment)

N;Alternate names: autoproteolysin IIA; plasma protein C

C;Species: Bos primigenius taurus (cattle)

C;Date: 30-Nov-1980 #sequence,revision 17-Mar-1987 #text_change 16-Jul-1999

C;Accession: A26250; A18385; A18386; A00928

R;Long, G.L.; Balagaje, R.M.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 81, 5653-5656, 1984

A;Title: Cloning and sequence of liver cDNA coding for bovine protein C.

A;Reference number: A26250; MUID:85014826

A;Accession: A26250

A:Molecule type: mRNA
 A:Residues: 1-456 <LON>
 R:Fernlund, P.; Stenflo, J.
 J. Biol. Chem. 257, 12170-12179, 1982
 A:Title: Amino acid sequence of the light chain of bovine protein C.
 A:Reference number: A18385; MUID:83007325
 A:Accession: A18385
 A:Molecule type: protein
 A:Residues: 40-194 <FER>
 A:Note: 82-Lys was also found
 R:Dirksenberg, T.; Roepstorff, P.; Stenflo, J.
 Proc. Natl. Acad. Sci. U.S.A. 80, 1802-1806, 1983
 A:Title: beta-hydroxyaspartic acid in vitamin K-dependent protein C.
 A:Reference number: A19316; MUID:83169769
 A:Contents: annotation; revision to residue 110
 R:Stenflo, J.; Fernlund, P.
 J. Biol. Chem. 257, 12180-12190, 1982
 A:Title: Amino acid sequence of the heavy chain of bovine protein C.
 A:Reference number: A18386; MUID:83007326
 A:Accession: A18386
 A:Molecule type: protein
 A:Residues: 197-454, 'PV' <STR>
 R:Esmon, N.L.; Debault, L.E.; Esmon, C.T.
 J. Biol. Chem. 258, 5548-5553, 1983
 A:Title: Proteolytic formation and properties of gamma-carboxyglutamic acid-domainless F
 A:Reference number: A37541; MUID:8321513
 A:Contents: annotation; activation; calcium binding
 R:Johnson, A.E.; Esmon, N.L.; Laue, T.M.; Esmon, C.T.
 J. Biol. Chem. 258, 5554-5560, 1983
 A:Title: Structural changes required for activation of protein C are induced by Ca2+ bin
 A:Reference number: A37542; MUID:8321514
 A:Contents: annotation; activation; calcium binding
 C:Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that re
 B:Comment: Protein C is synthesized in the liver as a single chain precursor, which is g
 bin, which cleaves a tetradecapeptide from the amino end of the heavy chain; this react
 C:Comment: Calcium binds to the gamma-carboxyglutamic acid (Gla) residues and, with strc
 cognition of the thrombin-thrombomodulin complex.
 C:Comment: The gamma-carboxyglutamic acid residues arise by a posttranslational, vitamin
 C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
 C:Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding
 F:1-29/Domain: signal sequence (fragment) #status predicted <SIG>
 F:24-83/Domain: Gla domain homology <Gla>
 F:30-39/Domain: propeptide #status predicted <PRO>
 F:40-194/Product: protein C light chain #status experimental <LCH>
 F:98-128/Domain: EGF homology <EG1>
 F:137-172/Domain: EGF homology <EG2>
 F:197-456/Product: protein C heavy chain #status experimental <HCH>
 F:197-210/Domain: activation peptide #status experimental <APR>
 F:211-440/Domain: trypsin homology <TRY>
 F:45, 46, 53, 55, 58, 59, 62, 64, 65, 68, 74/Modified site: gamma-carboxyglutamic acid (Glu) #stat
 F:110/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental
 F:119-128, 137-148, 144-157, 159-172, 180-318, 237-253, 368-382, 393-421/Dsulfide bonds: #stat
 F:136, 289, 350/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:252, 298, 337/Active site: His, Asp, Ser #status predicted
 F:366/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 64.5%; Score 129; DB 1; Length 456;
 Best Local Similarity 54.8%; Pred. NO. 2.7e-13;
 Matches 23; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 1 ANSFLXLRHGSLRXRCIXXICDFFXXAKXIFONVDDTLAFWS 42
 DB 40 ANSFLELRPGNVERCESECEFEAREITFONTEDTAFWS 81

RESULT 5
 EXRT
 coagulation factor Xa (EC 3.4.21.6) precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Jan-1995 #sequence_revision 07-Feb-1997 #text_change 08-Dec-2000
 C:Accession: S49075; J04670; PS0190; I62745

R:Stanton, C.; Ross, P.; Hutson, S.; Wallin, R.
 Thromb. Res. 80, 63-73, 1995
 A:Title: Evidence for competition between vitamin K-dependent clotting factors for in
 A:Reference number: A58498; MUID:96093366
 A:Accession: S49075
 A:Molecule type: mRNA
 A:Residues: 1-482 <STAL>
 A:Cross-references: EMBL:X79807; NID:9506600; PIDN:CA56202.1; PID:9506601
 A:Note: submitted to the EMBL Data Library, June 1994
 A:Note: neither the complete nucleic acid sequence nor the complete translation are s
 R:Stanton, C.; Ross, P.; Hutson, S.; Wallin, R.
 Gene 169, 269-273, 1996
 A:Title: Processing and expression of rat and human clotting factor-X-encoding cDNAs.
 A:Reference number: J04670; MUID:96194815
 A:Accession: J04670
 A:Molecule type: mRNA
 A:Residues: 1-482 <STAL>
 A:Cross-references: EMBL:X79807; NID:9506600; PIDN:CA56202.1; PID:9506601
 A:Experimental source: Cos-1 cell
 R:Enjoly, K.; Miyazaki, K.; Kato, H.
 J. Biochem. 109, 890-898, 1991
 A:Title: Characterization of rat factors X and Xa: demonstration of factor Xa in rat
 A:Reference number: PS0190; MUID:92041742
 A:Accession: PS0191
 A:Molecule type: protein
 A:Residues: 41-58, 'X', 60-65 <ENJ>
 A:Accession: PS0190
 A:Molecule type: protein
 A:Residues: 183-186, 'X', 188-207 <ENJ2>
 R:Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
 Eur. J. Haematol. 52, 162-168, 1994
 A:Title: Analysis of the partial nucleotide sequences and deduced primary structures
 A:Reference number: I46196; MUID:94222160
 A:Accession: I62745
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 295-383, 'G', 385-455 <MUR>
 A:Cross-references: GB:D21215; NID:9415309; PIDN:BA04756.1; PID:9455396
 C:Function:
 A:Description: catalyzes the proteolytic activation of prothrombin to thrombin in the
 A:Pathway: blood coagulation
 C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol
 C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-40/Domain: propeptide #status predicted <PRO>
 F:25-84/Domain: Gla domain homology <Gla>
 F:41-179/Product: coagulation factor X light chain #status predicted <LCH>
 F:90-121/Domain: EGF homology <EG1>
 F:129-164/Domain: EGF homology <EG2>
 F:183-482/Product: coagulation factor X heavy chain #status predicted <HCH>
 F:183-231/Domain: activation peptide #status predicted <APR>
 F:232-482/Product: coagulation factor Xa heavy chain #status predicted <ACT>
 F:232-460/Domain: trypsin homology <TRY>
 F:46, 47, 54, 56, 59, 60, 65, 66, 69, 72, 79/Modified site: gamma-carboxyglutamic acid (Glu) #s
 F:57-62, 90-101, 95-110, 112-121, 129-149, 136-149, 151-164, 172-340, 238-243, 259-275, 388-402
 F:103/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
 F:187/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:208/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F:218/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:231-232/Cleavage site: Arg-Tile (coagulation factor IXa, coagulation factor VIIa) #s
 F:274, 320, 417/Active site: His, Asp, Ser #status predicted

Query Match 52.5%; Score 105; DB 1; Length 482;
 Best Local Similarity 38.6%; Pred. NO. 2.6e-09;
 Matches 17; Conservative 11; Mismatches 16; Indels 0; Gaps 0;

QY 1 ANSFLXLRHGSLRXRCIXXICDFFXXAKXIFONVDDTLAFWSK 44
 DB 41 ANSFEEIKKGNLRECEVEICISFEAREVPEDNKTEFWMKY 84

RESULT 6

A:Accession: B42696
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 383-617, 'E' <BAN>
A:Cross-references: GB:M01397
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydrolysis
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-43/Domain: propeptide #status predicted <PRO>
F:28-88/Domain: Gla domain homology <GLA>
F:44-617/Product: prothrombin #status experimental <PMAT>
F:109-187/Domain: kringle homology <KR1>
F:215-292/Domain: kringle homology <KR2>
F:360-609/Domain: trypsin homology <TRY>
F:50-51,58-60,63-64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu) #status
F:61-66,91-104,109-187,130-170,158-182,215-292,236-276,264-287,332-478,387-403,532-546,5
F:403,458,564/Active site: His, Asp, Ser #status predicted

Query Match 41.2%; Score 82.5; DB 2; Length 617;
Best Local Similarity 40.0%; Pred. No. 1.7e-05;
Matches 18; Conservative 7; Mismatches 19; Indels 1; Gaps 1;

Oy 1 ANS-FLXXLRHGSIXKXCIXICDFXKXKXIFQNVDTLAFWSKH 44
Db 44 ANSGLEELRKGNLERECVECCSYEEAFEALESPODIDVFWAKY 88

RESULT 12

thrombin (EC 3.4.21.5) B chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 14-Dec-1990 #sequence revision 14-Dec-1990 #text change 22-Jun-1999
C:Accession: A35827; A42696; S12081
R:Deben, S.J.F.; Schaefer, L.A.; Jamison, C.S.; Grant, S.G.; Fitzgibbon, J.J.; Pal, J.A.
DNA Cell Biol. 9, 487-498, 1990
A:Title: Characterization of the cDNA coding for mouse prothrombin and localization of the
A:Reference number: A35827; MUID:91025551
A:Accession: A35827
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-618 <DEC>
A:Cross-references: GB:M52308; NID:G53813; PIDN:CAA36548.1; PID:G53814
A:Experimental source: strain C57BL/6
A:Note: the data were obtained from females resulting from the cross of M. domesticus and
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq
A:Reference number: A42696; MUID:92212913
A:Accession: A42696
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 384-618, 'E' <BAN>
A:Cross-references: GB:M01394
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: blood coagulation; calcium binding; carboxylglutamic acid; glycoprotein; hydrolysis
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-43/Domain: propeptide #status predicted <PRO>
F:28-88/Domain: Gla domain homology <GLA>
F:44-617/Product: prothrombin B #status predicted <MAT>
F:109-187/Domain: kringle homology <KR1>
F:215-293/Domain: kringle homology <KR2>
F:361-610/Domain: trypsin homology <TRY>
F:50-51,58-60,63-64,69,70,73,76/Modified site: gamma-carboxylglutamic acid (Glu) #status
F:61-66,91-104,109-187,130-170,158-182,215-293,236-276,264-288,333-479,388-404,533-547,5
F:403,459,565/Active site: His, Asp, Ser #status predicted

Query Match 41.2%; Score 82.5; DB 2; Length 618;
Best Local Similarity 40.0%; Pred. No. 1.7e-05;
Matches 18; Conservative 7; Mismatches 19; Indels 1; Gaps 1;

Oy 1 ANS-FLXXLRHGSIXKXCIXICDFXKXKXIFQNVDTLAFWSKH 44

Db 44 ANSGLEELRKGNLERECVECCSYEEAFEALESPODIDVFWAKY 88

RESULT 13

plasma protein S precursor, vitamin K dependent - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 19-Mar-1997 #sequence revision 18-Jul-1997 #text change 16-Jul-1999
C:Accession: S53434
R:Greengard, J.S.; Fernandez, J.A.; Radtke, K.P.; Griffin, J.H.
Biochem. J. 305, 397-403, 1995
A:Title: Identification of candidate residues for interaction of protein S with C4b
A:Reference number: S53434; MUID:95134217
A:Accession: S53434
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-642 <GRE>
A:Cross-references: EMBL:L31380
A:Experimental source: tissue type liver
A:Note: the source is designated as rhesus monkey
C:Genetics:
A:Gene: PROS
C:Superfamily: plasma protein S; EGF homology; Gla domain homology; laminin G repeat
F:1-51/Domain: Gla domain homology (fragment) <GLA>
F:1-7/Domain: signal sequence (fragment) #status predicted <SIG>
F:8-642/Product: plasma protein S #status predicted <MAT>
F:87-120/Domain: EGF homology <EG1>
F:127-165/Domain: EGF homology <EG2>
F:171-207/Domain: EGF homology <EG3>
F:213-248/Domain: EGF homology <EG4>
F:281-633/Domain: sex hormone-binding globulin homology <SHB>
F:291-444/Domain: laminin G repeat homology <LGR>

Query Match 41.0%; Score 82; DB 2; Length 642;
Best Local Similarity 36.4%; Pred. No. 2.1e-05;
Matches 16; Conservative 10; Mismatches 18; Indels 0; Gaps 0;

Oy 1 ANSFLXXLRHGSIXKXCIXICDFXKXKXIFQNVDTLAFWSKH 44
Db 8 ANSMLEETKQGNLERECIEELCKKEAREVFENDDETDFYPPKY 51

RESULT 14

plasma protein S precursor - human
N:Alternate names: vitamin K-dependent protein S
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1990 #sequence revision 26-Jan-1996 #text change 16-Jul-1999
C:Accession: A35610; A35611; A26157; A25891; A36124; S02424; S09519
R:Schmidel, D.K.; Tatro, A.V.; Phelps, L.G.; Tomczak, J.A.; Long, G.L.
Biochemistry 29, 7845-7852, 1990
A:Title: Organization of the human protein S genes.
A:Reference number: A35610; MUID:91084444
A:Accession: A35610
A:Molecule type: DNA
A:Residues: 1-676 <SCH>
A:Cross-references: GB:M57853; NID:G190547; PIDN:AAA60357.1; PID:G190549; GB:J02917
A:Note: the authors translated the codon TTT for residue 26 as Leu
R:Ploos van Amstel, H.K.; Reitsma, P.H.; van der Logt, C.P.E.; Bertina, R.M.
Biochemistry 29, 7853-7861, 1990
A:Title: Intron-exon organization of the active human protein S gene Psalpa and its
A:Reference number: A35611; MUID:91084445
A:Accession: A35611
A:Molecule type: DNA
A:Residues: 1-25 <PL3>
A:Cross-references: GB:J02918
R:Hoskins, J.; Norman, D.K.; Beckmann, R.J.; Long, G.L.
Proc. Natl. Acad. Sci. U.S.A. 84, 349-353, 1987
A:Title: Cloning and characterization of human liver cDNA encoding a protein S precursor
A:Reference number: A26157; MUID:87092407
A:Accession: A26157

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2001, 08:50:31 ; Search time 12.45 Seconds

(without alignments)
79,530 Million cell updates/sec

Title: US-09-497-591-1

Perfect score: 200
Sequence: 1 ANSFLXXLRHGSLSXRCIXX.....XXAKXIFQNVDDTLAFWSKH 44Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176	88.0	44	3	US-08-955-636-1
2	176	88.0	45	2	US-08-965-832-2
3	176	88.0	419	1	US-08-295-411-1
4	176	88.0	419	2	US-08-955-471-1
5	176	88.0	419	5	PCT-US92-10242-1
6	176	88.0	460	2	US-08-756-506-2
7	176	88.0	460	2	US-08-756-506-4
8	176	88.0	460	6	5270178-13
9	176	88.0	460	6	5270178-14
10	176	88.0	460	6	5270178-15
11	176	88.0	460	6	5270178-16
12	176	88.0	461	6	5255577-2
13	176	88.0	461	6	5270178-17
14	176	88.0	461	6	5270178-18
15	176	88.0	461	6	5460953-3
16	168	88.0	44	3	US-08-955-636-19
17	168	88.0	44	3	US-08-955-636-24
18	166	83.0	44	3	US-08-955-636-35
19	163	81.5	42	2	US-08-745-254A-2
20	163	81.5	461	6	5270178-2
21	160	80.0	44	3	US-08-955-636-20
22	160	80.0	44	3	US-08-955-636-21
23	160	80.0	44	3	US-08-955-636-25
24	159	79.5	41	1	US-08-229-280-5
25	159	79.5	44	3	US-08-955-636-22
26	145	72.5	410	4	US-09-065-872-1
27	137	68.5	409	4	US-09-065-872-2

28	133	66.5	44	3	US-08-955-636-23	Sequence 23, Appl
29	123	61.5	44	3	US-08-955-636-2	Sequence 2, Appl
30	107	53.5	139	1	US-08-330-978-2	Sequence 2, Appl
31	107	53.5	139	1	US-08-474-042-2	Sequence 2, Appl
32	107	53.5	139	1	US-08-484-558-2	Sequence 2, Appl
33	107	53.5	139	1	US-08-774-592-2	Sequence 2, Appl
34	107	53.5	437	1	US-08-487-037-2	Sequence 2, Appl
35	107	53.5	437	1	US-08-487-037-3	Sequence 3, Appl
36	107	53.5	488	1	US-08-487-037-1	Sequence 1, Appl
37	101	50.5	487	1	US-08-469-486-53	Sequence 53, Appl
38	101	50.5	487	2	US-08-469-658-53	Sequence 53, Appl
39	101	50.5	492	1	US-08-469-486-2	Sequence 2, Appl
40	101	50.5	492	2	US-08-469-658-2	Sequence 2, Appl
41	100	50.0	448	1	US-08-295-411-3	Sequence 3, Appl
42	100	50.0	448	2	US-08-955-471-3	Sequence 3, Appl
43	100	50.0	448	5	PCT-US92-10068-1	Sequence 1, Appl
44	100	50.0	448	5	PCT-US92-10242-3	Sequence 3, Appl
45	96	48.0	44	3	US-08-955-636-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-955-636-1
Sequence 1, Application US/08955636A
Patent No. 6017882
GENERAL INFORMATION:
APPLICANT: Nelsestuen, Gary
TITLE OF INVENTION: MODIFIED VITAMIN K-DEPENDENT
FILE REFERENCE: 09531/002001
CURRENT APPLICATION NUMBER: US/08/955,636A
CURRENT FILING DATE: 1997-10-23
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 44
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (0)...(0)
OTHER INFORMATION: Xaa-gamma carboxyglutamic acid or glutamic acid
US-08-955-636-1

Query Match 88.0%; Score 176; DB 3; Length 44;
Best Local Similarity 97.7%; Pred. No. 7.5e-23;
Matches 43; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGSLSXRCIXXIDFFXXAKXIFQNVDDTLAFWSKH 44
Db 1 ANSFLXXLRHGSLSXRCIXXIDFFXXAKXIFQNVDDTLAFWSKH 44

RESULT 2
US-08-965-832-2
Sequence 2, Application US/08965832
Patent No. 5847085
GENERAL INFORMATION:
APPLICANT: CHARLES T. ESMON AND MIKHAIL D. SMIRNOV
TITLE OF INVENTION: Modified Protein C
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentlin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,832
; FILING DATE: 7-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/745,254
; FILING DATE: 8-NOV-1996
; PRIOR APPLICATION DATA: 60/053,768
; APPLICATION NUMBER: 60/053,768
; FILING DATE: 25-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patricia L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR# 165/167
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; LOCATION:
; OTHER INFORMATION: /note= "partial sequence of human protein C"
;
US-08-965-832-2
;
Query Match      88.0%; Score 176; DB 2; Length 45;
Best Local Similarity 97.7%; Pred. No. 7.7e-23;
Matches 43; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ANSEFLXXLRH9SLXRXCIXXICDFXAXKXIFQNVDDTLAFWSKH 44
Db 1 ANSEFLXXLRHSSLXRXCIXXICDFXAXKXIFQNVDDTLAFWSKH 44
;
RESULT 3
US-08-295-411-1
; Sequence 1, Application US/08295411
; Patent No. 5679639
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Westers, Rolf M.
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentlin Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/295,411
; FILING DATE: 22-AUG-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,989
; FILING DATE: 18-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSRI263.0C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..157
; OTHER INFORMATION: /note= "Protein C Light Chain"
;
US-08-295-411-1
;
Query Match      88.0%; Score 176; DB 1; Length 419;
Best Local Similarity 77.3%; Pred. No. 9.3e-22;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 ANSEFLXXLRH9SLXRXCIXXICDFXAXKXIFQNVDDTLAFWSKH 44
Db 1 ANSEFLXXLRHSSLXRXCIXXICDFXAXKXIFQNVDDTLAFWSKH 44
;
RESULT 4
US-08-955-471-1
; Sequence 1, Application US/08955471
; Patent No. 5968751
; GENERAL INFORMATION:
; APPLICANT: Griffin, John H.
; APPLICANT: Westers, Rolf M.
; TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
; TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
; TITLE OF INVENTION: for Inhibiting Coagulation
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10666 No. 5968751th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentlin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

1 APPLICATION NUMBER: US/08/955,471
2 FILING DATE:
3 CLASSIFICATION:
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: 08/295,411
6 FILING DATE:
7 CLASSIFICATION:
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Fitting, Thomas
10 REGISTRATION NUMBER: 34,163
11 REFERENCE/DOCKET NUMBER: TSRI263, 0C1
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: 619-554-2937
14 TELEFAX: 619-554-6312
15 INFORMATION FOR SEQ ID NO: 1:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 419 amino acids
18 TYPE: amino acid
19 TOPOLOGY: linear
20 MOLECULE TYPE: protein
21 HYPOTHEICAL: NO
22 ANTI-SENSE: NO
23 FEATURE:
24 NAME/KEY: Region
25 LOCATION: 1..157
26 OTHER INFORMATION: /note= "Protein C Light Chain"
27 FEATURE:
28 NAME/KEY: Region
29 LOCATION: 158..169
30 OTHER INFORMATION: /note= "Protein C Activation
31 OTHER INFORMATION: Peptide"
32 FEATURE:
33 NAME/KEY: Region
34 LOCATION: 170..419
35 OTHER INFORMATION: /note= "Protein C Heavy Chain"
36
37 US-08-955-471-1
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
98
```

[illegible]

THIS PAGE BLANK (ustp10)

PT Vitamin K-dependent polypeptide with modified gamma-carboxyglutamic
PT acid domain, useful for treating clotting disorders
XX
PS Disclosure; Page 14; 86pp; English.
XX
CC This sequence is the protein C GLA (gamma-carboxyglutamic acid)
CC domain. The invention relates to a vitamin K-dependent polypeptide
CC comprising a modified GLA domain containing an amino acid substitution
CC which enhances membrane binding of the modified polypeptide as compared
CC to the native polypeptide. The polypeptide is used to treat a clotting
CC disorder by decreasing or increasing clot formation. Modification of the
CC GLA domain results in a protein which has enhanced membrane binding
CC affinity as compared to the native protein.
XX
SQ Sequence 44 AA:

Query Match 88.0%; Score 176; DB 20; Length 44;
Best Local Similarity 97.7%; Pred. No. 2e-22; Mismatches 0; Gaps 0;
Matches 43; Conservative 0; Indels 1;

OY 1 ANSFLXLRHGSIXRCIXICDFXXAKXIFQNVDDTLAFWSKH 44
Db 1 ansflxlrhgsixrcixicdfxxakxifgnvddtlafwskh 44
|||||

RESULT 2
AAB36402 standard; peptide; 44 AA.
XX
AC AAB36402:
XX
DT 27-FEB-2001 (first entry)
XX
DE Human protein C gamma-carboxyglutamic acid domain SEQ ID NO:1.
XX
KW Vitamin K-dependent protein; factor VII; protein C; GLA domain;
KW gamma-carboxyglutamic acid domain; factor IX; protein S; protein Z;
KW factor X; prothrombin; enhanced membrane binding affinity;
KW clot formation; thrombolytic; haemostatic; bleeding disorder;
KW thrombosis; clotting disorder; haemophilia A; haemophilia B;
KW liver disease.
XX
OS Homo sapiens.
XX
PN MO200066753-A2.
XX
PD 09-NOV-2000.
XX
PF 28-APR-2000; 2000WO-US11416.
XX
PR 29-APR-1999; 99US-0302239.
XX
PA (MINU) UNIV MINNESOTA.
XX
PI Nelsestuen GL;
XX
DR WPI; 2001-007226/01.
XX
PT Novel vitamin K-dependent polypeptide useful for treating clotting
PT disorders such as thrombosis and hemophilia, comprises modified
PT gamma-carboxy glutamic acid domain that enhances membrane binding
PT affinity -
XX
PS Example 5; Page 42; 81pp; English.
XX
CC The present invention describes a vitamin K-dependent polypeptide (I)
CC comprising a modified gamma-carboxy glutamic acid (GLA) domain having
CC at least one amino acid substitution, that enhances membrane binding
CC affinity and the activity of the polypeptide relative to a corresponding
CC native vitamin K-dependent polypeptide and inhibits clot formation.
CC (I) can have thrombolytic and haemostatic activities, and can be used
CC as an inhibitor of clot formation. (I) is useful for decreasing clot

CC formation in a mammal, a factor VII or factor IX containing a modified
CC GLA domain is useful for increasing clot formation and for treating a
CC bleeding disorder, including thrombosis and clotting disorders such as
CC haemophilia A, haemophilia B and liver disease. The present sequence
CC represents a human protein C GLA domain sequence, given in the
CC exemplification of the present invention.
XX
SQ Sequence 44 AA:

Query Match 88.0%; Score 176; DB 22; Length 44;
Best Local Similarity 97.7%; Pred. No. 2e-22; Mismatches 0; Gaps 0;
Matches 43; Conservative 0; Indels 1;

OY 1 ANSFLXLRHGSIXRCIXICDFXXAKXIFQNVDDTLAFWSKH 44
Db 1 ansflxlrhgsixrcixicdfxxakxifgnvddtlafwskh 44
|||||

RESULT 3
AAW5710 standard; protein; 45 AA.
XX
AC AAW5710:
XX
DT 08-DEC-1998 (first entry)
XX
DE Partial human protein C amino acid sequence.
XX
KW Gamma carboxyglutamic acid; human protein C; GLA domain; chimera;
KW pRC/RSV; RSV-PC; amplification, PCR, primer; transfection; anticoagulant;
KW human 293 cell; Protein S; myocardial infarction; venous thrombosis;
KW disseminated intravascular coagulation; thromboembolic disease; lupus;
KW adult respiratory distress syndrome; factor V Leiden; stroke.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Misc-difference 6 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 7 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 14 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 16 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 16 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 19 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 20 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 25 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 25 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 26 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 29 /note= "Gamma carboxyglutamic acid"
FT Misc-difference 29 /note= "Gamma carboxyglutamic acid"
XX
XX WO9820118-A1.
XX
PD 14-MAY-1998.
XX
XX 07-NOV-1997; 97WO-US20376.
XX
PF 25-JUL-1997; 97US-0053768.
XX
PR 08-NOV-1996; 96US-0745254.
XX
PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
XX
PI Esmon CT, Smirnov M;
XX
DR WPI; 1998-286934/25.
XX
PT Protein C chimeric proteins for use as anticoagulants - having gamma


```

xx      The PC polypeptides indicated in the Features Table inhibit
CC      coagulation (they prevent binding of serine protease to natural
CC      substrates), esp. when admin. to give an Intravascular blood
CC      concn. of 0.1-100 (pref. 0.5-10) microm.
CC      NB: Sequences corresp. to SEQ ID NO 6, 7, 8 and 9 are described
CC      in the specification but have not yet been added to the SEQUENCE
CC      LISTING.
xx
xx      Sequence      419 AA;
SQ
Query Match      88.0%;      Score 176;      DB 14;      Length 419;
Best Local Similarity 77.3%;      Pred. No. 2.2e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Oy      1 ANSFLXLRHGSLXRKCTIXICDFXKXKIRPNVDLTAFMSKSH 44
      ||||| ||| | | | | ||| | | -||| ||||| ||||| |||
db      1 ANSleeIRHSLrecleelcdfeakellrpnvdllafwskh 44

```

RESULT	6
AAW72753	
ID	AAW72753 standard; protein; 419 AA.

AC AAW72753;

DT 08-JAN-1999 (first entry)

DE Primary structure of activated human protein C

KW Human; activated protein C; primary structure; autodegradation;
 KW purification; processing; intravascular coagulation; thrombotic stroke;
 KW deep vein thrombosis; pulmonary embolism; peripheral arterial thrombosis
 KW emboli; heart; peripheral artery; acute myocardial infarction;
 KW disseminated intravascular coagulation; acute precapillary occlusion;
 KW precapillary occlusion; transplantation; retina thrombosis.
 OS Homo sapiens.
 XX

FT	Key	Location/Qualifiers
FT	Disulfide-bond	17..22
FT	Disulfide-bond	50..65
FT	Disulfide-bond	80..85
FT	Disulfide-bond	98..109
FT	Disulfide-bond	105..118
FT	Disulfide-bond	120..133
FT	Disulfide-bond	141..277
FT	Misc-difference	151

FT	Misc-difference	154	/note=	"unspecified"
FT	Misc-difference	155	/note=	"unspecified"
FT	Misc-difference	156	/note=	"unspecified"
FT	Misc-difference	157	/note=	"unspecified"
FT	Misc-difference	158	/note=	"unspecified"
FT	Misc-difference	159	/note=	"unspecified"
FT	Misc-difference	160	/note=	"unspecified"
FT	Misc-difference	161	/note=	"unspecified"
FT	Misc-difference	162	/note=	"unspecified"
FT	Misc-difference	163	/note=	"unspecified"
FT	Misc-difference	164	/note=	"unspecified"
FT	Misc-difference	165	/note=	"unspecified"

FT		/note=	"unspecified"
FT	Misc-difference	166	
FT		/note=	"unspecified"
FT	Misc-difference	167	
FT		/note=	"unspecified"
FT	Misc-difference	168	
FT		/note=	"unspecified"
FT	Misc-difference	169	
FT		/note=	"unspecified"
FT	Disulfide-bond	196..212	
FT	Disulfide-bond	351..345	
FT	Disulfide-bond	351...384	

XX	EP875563-A2.
PN	
XX	
PD	04-NOV-1998.

PF 28-APR-1998;

PR 28-APR-1997; 97US-0045255

PA (ELL) LILLY & CO ELL.

PI Baker JC, Carlson AD, Huang L, Shelliga TA;

DR WPI; 1998-559430/48.

Improved processing an

PT and pH to reduce the amount of autodegradation products

PS Disclosure; Fig 1; 7pp; English.

CC An improved method has been developed of processing an aqueous solution
CC of activated protein C (aPC). The method comprises conducting the
CC processing at an ionic strength greater than 150 mM and a pH of 5.5-6.3
CC The present sequence represents the primary structure of activated human
CC protein C, used to assist in illustrating the autodegradation pathways
CC described in the present invention. The aPC is used for treating a
CC variety of acquired disease states involving intravascular coagulation
CC e.g. thrombotic stroke, deep vein thrombosis, pulmonary embolism,
CC peripheral arterial thrombosis, emboli originating from the heart or
CC peripheral arteries, acute myocardial infarction, disseminated
CC intravascular coagulation and acute pre- or postcapillary occlusions
CC including transplantations or retina thrombosis. The process minimises
CC the percentage of autodegradation products forming at most 10%,
CC preferably below 5% of des(1-9) aPC and des(1-10) aPC by weight, to
CC achieve a more potent, higher purity aPC pharmaceutical preparation.

50 Sequence 419 AA

Query Match	88.0%	Score 176;	DB 19;	Length 419;
Best Local Similarity	77.3%;	Pred. No. 2.2e-21;		
Matches 34;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;

```
QY      1 ANSFLXLLRHGSLXRXCIXICDFXXAKXIIFQVNDDTLAFWSKH 44
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 ansfleelrhhslerecieicdfeeakeifqnvdttlafwsh 44
```

RESULT	7
AAB36894	
ID	AAB36894 standard; Protein; 419 AA

AC AAB36894

DT 26-FEB-2001 (first entry)

DE Human protein C derivative 1.

KW Protein C; human; vascular occlusive; burn; transplantation;

PF	13-APR-2000; 2000MO-US08722.
XX	
PR	30-APR-1999; 99US-0131801.
XX	
XX	(EELI) LILLY & CO ELI.
PA	
XX	
XX	
PI	Gerlitz BE, Jones BE;
XX	
DR	WPI; 2001-007227/01.
DR	N-PSDB; AAC83313.
PT	Protein C derivatives, useful for treating vascular occlusive disorder
PT	hypercoagulable state, thrombotic disorder and disease states
PT	predisposing thrombosis, comprises specific amino acid substitutions -
XX	
XX	
PS	Claim 4: Page 46-47; 57pp; English.
XX	
CC	The present invention relates to a human protein C derivative. The
CC	protein is useful for treating vascular occlusive disorders,
CC	hypercoagulable states such as sepsis, disseminated intravascular
CC	coagulation, purpura fulminans, major trauma, major surgery, burns,
CC	adult respiratory distress syndrome, transplantation, deep vein
CC	thrombosis, heparin-induced thrombocytopenia, sickle cell disease,
CC	thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic

CC The present invention relates to a human protein C derivative. The
CC protein is useful for treating vascular occlusive disorders,
CC hypercoagulable states such as sepsis, disseminated intravascular
CC coagulation, purpura fulminans, major trauma, major surgery, burns,
CC adult respiratory distress syndrome, transplantation, deep vein
CC thrombosis, heparin-induced thrombocytopenia, sickle cell disease,
CC thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic
CC purpura, and hemolytic uremic syndrome, and also useful for treating
CC thrombotic disorders and acute coronary syndromes such as myocardial
CC infarction, unstable angina, and stroke. Protein C derivatives with
CC amino acid substitutions result in increased resistance to
CC inactivation by serpins when compared to wild-type activated human
CC protein C. They also have longer half-lives in human blood and hencece
CC require either less frequent administration and/or smaller dosage
CC than wild type human protein C for treating disorders.
XX

SQ Sequence 419 AA:

Query Match 88.0%; Score 176; DB 22; Length 419;
Best Local Similarity 77.3%; Pred. No. 2,2e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps

	Query Match	88.0%	Score 176;	DB 22;	Length 419;
	Best Local Similarity	77.3%	Pred. No.	2.2e-21;	
	Matches 34; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;
Oy	1 ANSFLXLRHGSLSKRXCIXXICDFYXAAXKIIFONVDTLAFAWSKH	44			
Dd	1 ansflxelrhtsslsercleeelcdfeeakellfgnvddcltafwskh	44			

```

Oy      1  ANSFXXLRHGS LKRXCIIXICDPYXAKKXIFQNVNDPFLAFMSKH 44
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      1  ansfleeirhs lerectleeicdtfeekelfgnvdtllafrskh 44

RESULT  9
AAB36897
ID      AAB36897 standard; Protein; 419 AA.
XX
XX      AAB36897;
XX

```

DE	Human protein C derivative 4.
XX	
KW	Protein C; human; vascular occlusive; burn; transplantation;
KW	deep vein thrombosis; sickle cell; thalassemia;
KW	thrombotic disorders; myocardial infarction; angina; stroke
XX	
OS	Homo sapiens.
XX	
PN	MO200066754-A1.
XX	
PD	09-NOV-2000.
XX	
PE	13-APR-2000; 2000MO-US08722.
XX	
FR	30-APR-1999; 99US-0131801.
XX	
PA	(ELIL) LILLY & CO ELI.
XX	
PI	Gerltz BE, Jones BE;
XX	

DR WPI: 2001-007227/01.
DR N-PSDB: AAC83314.
XX Protein C derivatives, useful for treating vascular occlusive disorder,
PT hypercoagulable state, thrombotic disorder and disease states
PT predisposing thrombosis, comprises specific amino acid substitutions -
XX
PS Claim 5: Page 48-49; 57pp: English.
XX
XX The present invention relates to a human protein C derivative. The
CC protein is useful for treating vascular occlusive disorders,
CC hypercoagulable states such as sepsis, disseminated intravascular
CC coagulation, purpura fulminans, major trauma, major surgery, burns,
CC adult respiratory distress syndrome, transplantation, deep vein
CC thrombosis, heparin-induced thrombocytopenia, sickle cell disease,
CC thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic
CC purpura, and hemolytic uremic syndrome, and also useful for treating
CC thrombotic disorders and acute coronary syndromes such as myocardial
CC infarction, unstable angina, and stroke. Protein C derivatives with
CC amino acid substitutions result in increased resistance to
CC inactivation by serpins when compared to wild-type activated human
CC protein C. They also have longer half-lives in human blood and hence
CC require either less frequent administration and/or smaller dosage
CC than wild type human protein C for treating disorders.
XX
SQ Sequence 419 AA:

Query Match 88.0%; Score 176; DB 22; Length 419;
Best Local Similarity 77.3%; Pred. No. 2.2e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGSIXKXICDFXXAKXIFQNVDDTLAFWSKH 44
Db 1 ansflxxlrhgsixkxicdfxxakxifqnvddtlaafwskh 44

RESULT 10
AAB36898
ID AAB36898 standard; Protein: 419 AA.
XX
XX AAB36898;
AC
XX 26-FEB-2001 (first entry)
DT
XX
XX Human protein C derivative 5.
DE
XX
XX Protein C: human; vascular occlusive; burn; transplantation;
KW deep vein thrombosis; sickle cell; thalassemia;
KW thrombotic disorders; myocardial infarction; angina; stroke.
XX
XX Homo sapiens.
OS
XX
XX WO20006754-A1.
PN
XX
XX 09-NOV-2000.
PD
XX
XX 13-APR-2000; 2000WO-US08722.
PF
XX
XX 30-APR-1999; 99US-0131801.
PR
XX
XX (ELIL) LILLY & CO ELI.
PA
XX
XX Gerlitz BE, Jones BE;
PI
XX
XX WPI: 2001-007227/01.
DR
XX
XX N-PSDB: AAC83315.
PT
XX Protein C derivatives, useful for treating vascular occlusive disorder,
PT hypercoagulable state, thrombotic disorder and disease states
PT predisposing thrombosis, comprises specific amino acid substitutions -
XX
XX Claim 6; Page 49-51; 57pp: English.

XX
CC The present invention relates to a human protein C derivative. The
CC protein is useful for treating vascular occlusive disorders,
CC hypercoagulable states such as sepsis, disseminated intravascular
CC coagulation, purpura fulminans, major trauma, major surgery, burns,
CC adult respiratory distress syndrome, transplantation, deep vein
CC thrombosis, heparin-induced thrombocytopenia, sickle cell disease,
CC thalassemia, viral hemorrhagic fever, thrombotic thrombocytopenic
CC purpura, and hemolytic uremic syndrome, and also useful for treating
CC thrombotic disorders and acute coronary syndromes such as myocardial
CC infarction, unstable angina, and stroke. Protein C derivatives with
CC amino acid substitutions result in increased resistance to
CC inactivation by serpins when compared to wild-type activated human
CC protein C. They also have longer half-lives in human blood and hence
CC require either less frequent administration and/or smaller dosage
CC than wild type human protein C for treating disorders.
XX
SQ Sequence 419 AA:

Query Match 88.0%; Score 176; DB 22; Length 419;
Best Local Similarity 77.3%; Pred. No. 2.2e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGSIXKXICDFXXAKXIFQNVDDTLAFWSKH 44
Db 1 ansflxxlrhgsixkxicdfxxakxifqnvddtlaafwskh 44

RESULT 11
AAB81104
ID AAB81104 standard; Protein: 460 AA.
XX
XX AAB81104;
AC
XX 16-SEP-1990 (first entry)
DT
XX
XX Sequence of human protein C.
DE
XX
XX Human protein C; plasmid pPC 1.
KW
XX
XX Homo sapiens.
OS
XX
XX JP63263083-A.
PN
XX
XX 30-NOV-1988.
PD
XX
XX 21-APR-1987; 87JP-0096341.
PF
XX
XX 21-APR-1987; 87JP-0096341.
PR
XX
XX (FARH) HOECHST JAPAN KK.
PA
XX
XX WPI: 1988-350711/49.
DR
XX
XX N-PSDB: AAN81408.
DT
XX
XX Human protein C gene - prepd. from new DNA of specified base sequence.
PT
XX
XX Disclosure; : 16pp; Japanese.
PS
XX
XX The human protein C is expressed in large ams. using plasmid pPC 1 in
CC E.coli K12/On 225 (FERM P-9297).
CC
XX
SQ Sequence 460 AA:

Query Match 88.0%; Score 176; DB 9; Length 460;
Best Local Similarity 77.3%; Pred. No. 2.4e-21;
Matches 34; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ANSFLXXLRHGSIXKXICDFXXAKXIFQNVDDTLAFWSKH 44
Db 42 ansflxxlrhgsixkxicdfxxakxifqnvddtlaafwskh 85

